CENTER FOR GLOBAL DEVELOPMENT

Advancing the Measurement of Violence in and Around Schools: Evidence from Malawi

Gabriela Smarrelli, Esme Kadzamira, Thi Le, Tionge Saka

Abstract

Approximately one billion children worldwide report having experienced physical, emotional, or sexual violence, yet existing evidence suggests this figure substantially underestimates true prevalence, as individuals often avoid disclosing their experiences of violence. We study which survey methods yield more reliable measures of violence against children. We randomly assigned 6,000 children aged 8 to 12 years across 251 schools in Malawi to respond to a survey using face-to-face (F2F) interviews or audio computer-assisted self-interviews (ACASI), administered either at the child's home or school. We find that ACASI, relative to F2F, substantially increases the disclosure of physical and emotional violence **by** peers and school staff, regardless of survey location. Disclosure of sexual violence more than doubles when using ACASI and when surveys are conducted at school rather than the child's home. These effects are consistent with the greater privacy provided by ACASI and the additional reassurance offered by school settings, which provide physical spaces that may more strongly signal to respondents that their answers are unlikely to be overheard, reducing social and emotional barriers to disclosure. Our results are robust to multiple checks examining social desirability bias, lapses in attention, and misunderstanding of survey questions, and suggest that ACASI offers a cost-effective approach to reducing systematic measurement error in surveys on sensitive topics.

KEYWORDS
measurement, survey methods, survey experiment, school-based violence, sexual violence, economics of education

JEL CODES
C81, C82, C83, I29, J13

Advancing the Measurement of Violence in and Around Schools: Evidence from Malawi

Gabriela Smarrelli
Center for Global Development
gsmarrelli@cgdev.org

Esme Kadzamira
University of Malawi
ekadzamira@unima.ac.mw

Thi Le
Center for Global Development
tle@cgdev.org

Tionge Saka
University of Malawi
twsaka@unima.ac.mw

Gabriela Smarrelli, Esme Kadzamira, Thi Le, Tionge Saka. 2026. "Advancing the Measurement of Violence in and Around Schools: Evidence from Malawi." CGD Working Paper 751. Washington, DC: Center for Global Development.
https://www.cgdev.org/publication/advancing-measurement-violence-and-around-schools-evidence-malawi

For the consent form and survey instruments used in this study, see the supplementary materials: https://www.cgdev.org/sites/default/files/supplementary-materials-advancing-measurement-violence-around-schools-malawi.pdf

We are especially grateful to Juliet Magombo for her excellent management and supervision of field activities. We are also grateful to Sofia Amaral, Lelys Dinarte, Oscar Diaz Botia, David Evans, Marta Favara, Markus Goldstein, Dipak Naker, Eeshani Kandpal, Susannah Hares, Radhika Nagesh, Alan Sanchez, and Justin Sandefur, as well as the participants at research seminars and conferences, for their comments, feedback and stimulating discussions. Many thanks as well to the Malawi Ministry of Education for supporting this research, and special thanks to YONECO, as well as the enumerators and counsellors who participated in the data collection process. We would also like to thank the Echidna Giving and Pousaz Philanthropies for the financial support. Our gratitude extends to the unmentioned individuals who supported us throughout this research project. This research received ethical approval from the University of Malawi Institutional Review Board (Ref No: P.04/24/347).

CENTER FOR GLOBAL DEVELOPMENT

2055 L Street, NW Fifth Floor
Washington, DC 20036
202.416.4000

1 Abbey Gardens
Great College Street
London
SW1P 3SE

www.cgdev.org

1 Introduction

Measuring sensitive issues through surveys poses major challenges. Topics such as sexual and physical violence, crime, drug use, and female genital cutting are difficult to measure accurately, as respondents may misreport either by concealing stigmatised experiences or providing socially desirable answers. This has negative repercussions for research, policy response and safeguarding. Misreporting not only leads to inaccurate statistics on the prevalence of harmful practices and abuse, but it can introduce systematic error and bias conclusions regarding the effectiveness of prevention interventions and policies. Misreporting violence in surveys – particularly underreporting – can also be a missed opportunity to identify cases and refer them to the necessary services and authorities to protect potential victims.

Researchers have long been interested in and studied how and which methods may reduce misreporting or limit bias in self-reported data on violence. However, the focus has been on adults, and particularly on intimate partner violence (Joseph et al. 2017; Bulte and Lensink 2019; Agüero and Frisancho 2022; Cullen 2023; Peterman et al. 2025). Less is known regarding how to collect reliable data on children's experiences of violence (Tanton et al. 2023), a context which requires even more care to ensure children understand the questions being asked and are comfortable and safe disclosing their experiences of violence.

In this paper, we aim to address this gap. Existing surveys estimate that one billion children worldwide experience violence (Hillis et al. 2016). Much of this violence occurs in and around schools: a third of adolescents worldwide experience bullying (UNESCO 2019), and, in some countries, up to 25 percent of sexual violence cases occur in schools (Smarrelli and Wong 2025). These estimates are based on nationally representative student self-reported data from surveys that have been collected using different methods and in different locations. Specifically, of the nine international surveys[1] that ask at least one question on bullying, corporal punishment, or sexual violence, three collect the data using face-to-face (F2F) interviews at the respondent's home, and six use self-administered questionnaires at the respondent's school. Researchers also use different data collection methods: a systematic review of school-based violence prevention interventions shows that 90 percent of studies use self-administered questionnaires, while the rest use F2F surveys (Evans, Smarrelli and Wong, forthcoming). Yet little is known about which combinations of survey methods used with children produce violence estimates with the

[1] See Figure A.1 in Appendix A for the list of surveys.

lowest measurement error (Tanton et al. 2023). Even less is known about how to collect reliable data from children below the age of 12, for whom self-administered questionnaires are not a feasible option due to concerns about reading ability (Evans et al. 2025).

Motivated by this, we study which survey methods minimise misreporting and yield more reliable measures of school-related violence among children aged 8 to 12. We do this through a large-scale randomized survey experiment in Malawi, where decade-old data indicate that bullying and sexual violence within the school community are common. We randomly assigned over 6,000 primary school children (across 251 schools in eight districts in the north, center, and south of Malawi) to one of four treatment arms, which varied in terms of the survey mode used and the location of the survey. Survey mode was either F2F or Audio Computer-Assisted Self-Interviews (ACASI), in which children listened to pre-recorded questions through headphones and entered their answers on a tablet. We prioritise ACASI over pen-and-paper or computer-based self-administered surveys to address concerns about children's reading comprehension levels. Survey location was either at the child's home or school, the two most common places where surveys are administered. Importantly, the survey was administered in three different local languages, and comprehensive safeguarding protocols were carefully implemented, including counselling and referral pathways, to protect children and minimize risks of harm.

Through these treatment arms, we exogenously vary the degree of privacy in the survey setting, a key determinant of disclosure, as perceived risks of privacy breaches may exacerbate fears of retaliation, judgment, or social exclusion and thereby influence reporting decisions. ACASI provides greater privacy than face-to-face interviews, as children record responses on a tablet without interacting with the enumerator. Survey location is also likely to affect disclosure, depending on whether children perceive different levels of privacy at their home or at school. In the analysis, we use those children allocated to respond via face-to-face interviews at home as the control group, reflecting the status quo in major international surveys such as the Violence Against Children Survey (VACS) and the Demographic and Health Survey (DHS).

We find that ACASI substantially increases disclosure across all forms of violence. Disclosure of physical bullying from peers and corporal punishment by school staff rises by 20 percent, emotional bullying by peers by 10 percent, and emotional violence by school staff by 88 percent. The effect is largest for sexual violence, which includes different forms of violence with or without physical contact, from sexual comments to forced sexual activity. In the control group, 14 percent reported experiencing any form of sexual violence; this more than doubles, to nearly

31 percent, under ACASI, mainly driven by violence perpetrated in and around schools by adults.

Survey location affects only the reporting of sexual violence, with children disclosing higher levels of violence when interviewed at school rather than at home. Differences between home and school settings matter to the extent that the physical space, the physical distance to the perpetrator, and the people typically present in the environment (though not during the interview itself) shape children's perceptions of privacy. Our results suggest that the physical distance to the perpetrator matters less, relative to conducting the survey in physical spaces that clearly signal privacy, specifically, spaces where the respondent is clearly outside the listening range of others.

We also examine whether the effect of survey treatment on children's disclosure of school-related violence varies by individual, family, and school characteristics. We find that the child's age is significantly negatively correlated with the likelihood of reporting physical violence only in the F2F arm, suggesting that older children are more likely to underreport physical violence when asked directly by an enumerator. We also find that, relative to face-to-face interviews, ACASI increases the likelihood that boys report sexual violence by 8 percentage points compared with girls. This suggests systematic misreporting by gender, where boys are more likely to disclose their experiences of violence when more privacy is provided.

These results are based on a sample designed to be nationally representative. The eight districts included in this study were selected using probability-proportional-to-size sampling from the 19 of Malawi's 28 districts in which the non-governmental organisation responsible for assessing and following up reports of violence operated. Comparisons of observable characteristics suggest that these 19 districts are broadly similar to the remaining districts in the country, increasing the likelihood that our results are broadly representative of Malawi as a whole.

We run a series of robustness checks to assess the validity of our results. We test whether social desirability bias is different by survey method, using the Children's Social Desirability Short (CSD-S) scale. We find that children in the ACASI arm are seven percentage points less likely to give socially desirable responses, suggesting that the additional privacy provided by ACASI gives children fewer incentives to misreport and hence systematic misreporting is less likely. However, we also find that children were more likely to misunderstand some of the questions about sexual violence or fail attention checks in the ACASI arm. We argue that the measurement error resulting from this is more likely to be random error rather than reflecting

incentives to give misleading responses. We account for these errors in multiple ways. First, we create outcomes for sexual violence that correct for any cases where the child misunderstood the question and did not experience violence. Second, acknowledging that the approach we follow has limitations, we drop from the sample the observations that failed the attention checks. We find that our estimates are robust to these and other checks, validating that ACASI is more likely to reduce misreporting (mainly underreporting) than F2F.

Our detailed cost data, which are typically lacking in the literature (Tanton et al. 2023), also show that the ACASI method is a cost-effective alternative. We find that ACASI costs USD 1 more per child, but it is more effective in eliciting higher reporting of violence (which we assume to be truthful). Specifically, we find that increasing the probability of reporting sexual violence by one percentage point costs USD 3.67 for ACASI at school, USD 5.10 for ACASI at home, and USD 7.69 for face-to-face at school.

Our results contribute to the literature on survey methodology and the measurement of sensitive topics. Researchers have studied which tools and methods can potentially limit bias in self-reported data, focusing mainly on the adult population (Tourangeau and Yan 2007; Gibson et al. 2015; Blattman et al. 2016; De Weerdt et al. 2020; Blair et al. 2020). The variety of survey methods tested includes direct questioning (e.g., F2F and ACASI) and indirect questioning, such as list experiments and random response techniques. The challenges in all these methods lie in determining which method will reliably provide a more credible measure of violence.

In the context of violence measurement, most of the existing evidence focuses on criminal behaviour and intimate partner violence (IPV). Cullen (2023), Gilligan et al. (2025), and Agüero and Frisancho (2022) compare face-to-face surveys with list experiments, with only Cullen (2023) finding that list experiments increase the disclosure of IPV. Other methods, such as random response techniques, have been found to improve disclosure relative to face-to-face surveys. In Bangladesh, for example, women report higher levels of sexual harassment when asked using random response techniques compared with face-to-face surveys (Boudreau et al. 2023). These indirect questioning methods are designed to offer greater confidentiality to respondents. However, the mixed results, along with evidence that these methods often confuse adult respondents (Kramon and Weghorst 2019), suggest that they are unlikely to be effective with younger populations.

Our results add to the few studies using ACASI approaches to explore its effects on violence disclosure. Cullen (2023) finds no difference between face-to-face and ACASI in Nigeria, whereas

Peterman et al. (2025) report higher disclosure of IPV with ACASI in Senegal. Evidence from a survey experiment in Ethiopia similarly shows that ACASI leads to higher disclosure of experiences of violent conflict, relative to face-to-face (von Russdorf et al. 2024). We show that, in a population of young children, ACASI elicits higher disclosure of school-related forms of violence relative to face-to-face surveys, likely driven by the greater privacy afforded by ACASI. A similar result is reported by Punjabi et al. (2021), who use a survey experiment to compare reports of school-related violence collected through ACASI and F2F methods among 1,000 primary-school children in Uganda. Consistent with our findings, the study finds that disclosure of sexual violence nearly doubled among children assigned to ACASI, although it does not find statistically significant differences in reports of bullying or corporal punishment. These findings suggest that increasing privacy during data collection can substantially improve the disclosure of sexual violence among school-age children across different contexts.

Only a handful of papers have focused on studying the measurement of school-related violence against children. In addition to Punjabi et al. (2021), Tanton et al. (2023) identify four studies in high-, middle-, and low-income countries. Among these, three focus on bullying and study how variations in face-to-face administration (e.g., changing who administers the survey) affect disclosure. Barr et al. (2017) is closest to our approach as they compare face-to-face surveys with a method that provides greater privacy: the sealed-envelope approach. The authors show that the latter yields higher reports of sexual violence, providing evidence on the importance of greater confidentiality. A limitation of their approach, however, is that all children responded to the survey using both methods, constraining the possibility of fully disentangling the effect of the survey mode, as responses with one method may influence responses with the other. We circumvent this constraint by randomly allocating children to one of four data-collection approaches.

Finally, to our knowledge, this paper is the first to analyze whether survey location affects the disclosure of violence. Surveys with questions on violence are generally administered at the respondent's home or school. For example, VACS and DHS, which ask about sexual violence, are collected at the respondent's home; while surveys asking about bullying, such as the Program for International Student Assessment (PISA), are typically administered in schools (Evans et al. 2025). Our findings show that location matters when asking about sexual violence, and suggest that administering surveys at the location of the incident or where the perpetrator is part of the same institution (e.g., teachers in the school) matters less if the interview is conducted in

a space that is clearly outside the listening range of others. Such settings can reduce fears of being overheard by the perpetrator and alleviate feelings of shame, embarrassment, or concern about others (family members or peers) listening, thereby addressing barriers that lead to misreporting.

The remainder of the paper proceeds as follows. Section 2 describes the context and experimental design. Section 3 outlines our empirical strategy. Section 4 presents the results on the disclosure of violence. Section 5 explores the correlates of misreporting, and Section 6 discusses the validity of our findings. Section 7 presents our cost-effectiveness analysis, and Section 8 concludes.

2 Context, experimental design and data

2.1 Context

Most data on school-related violence comes from international surveys that include questions on at least one form of violence (bullying from peers, corporal punishment from teachers, and sexual violence). In Malawi, two such surveys, collected more than a decade ago, revealed a high incidence of school-related violence in the country. The Global School Health Surveys (GSHS), collected using self-administered questionnaires in the students' school, show that in 2009, 45 percent of children aged 11 to 16 years self-reported experiencing bullying from peers. This is 7 percentage points higher than in other low-and middle-income countries within a window of +/- 3 years (Smarrelli and Wong 2025). The VACS, administered using face-to-face methods in the respondent's home, estimates that 22 percent of girls and 17 percent of boys aged 13 to 24 years reported experiencing sexual violence over their lives. Of these cases, 28 percent and 22 percent, respectively, occurred in the school setting.

As in Malawi, 61 percent of low- and middle-income countries and 72 percent of high-income countries rely on self-administered school-based questionnaires to measure the prevalence of bullying, while 45 percent of low- and middle-income countries use home-based, face-to-face surveys to measure the prevalence of sexual violence perpetrated by peers or school staff.[2] From this, we learn that the two main survey methods used to collect data at scale are face-to-face surveys and self-administered questionnaires (where children read and answer the questions directly), and that these surveys are typically administered either at home or in school. A

[2]The estimates are based on a dataset from Evans, Hares, Smarrelli, and Wu (2025). High-income countries currently do not have comparable surveys collecting data on sexual violence against children.

systematic review about school-based violence prevention interventions further shows that 90 percent of studies used self-administered questionnaires, while the rest used F2F surveys (Evans, Smarrelli and Wong, forthcoming). However, we do not know which combination of survey method and location would potentially provide more reliable measures of school-related violence.

2.2 Experimental design

When collecting data on sensitive issues such as violence, there is a risk that it will be measured with some degree of error. Adults and children may have incentives to misreport their experiences of violence, often due to fears of breaches of privacy and confidentiality that could lead to retaliation, exclusion, or judgment. Moreover, even without incentives to misreport, data can still be inaccurate due to errors related to misunderstanding survey questions or lapses in attention. Minimizing these sources of error is crucial for obtaining an accurate understanding of the prevalence of violence, and, importantly, for producing unbiased estimates of the impacts of interventions, policies or laws.

In the context of program evaluations, we worry the most about systematic reporting errors that occur when a respondent intentionally conceals the truth. In such cases, if the measurement error in violence outcomes is correlated with the treatment or intervention being assessed, estimated treatment effects will potentially be biased, affecting our understanding of what works to prevent violence. We would worry less about random errors, which would reduce the precision of our estimates but would not introduce bias.

We design a survey experiment to identify survey practices that reduce the likelihood of misreporting of school-related violence, or in other words, the practices that would reduce the risk of systematic reporting errors. In selecting which practices to test, we considered three key criteria: (i) ensuring that the design allowed us to vary privacy levels; (ii) prioritising methods that could be used in large-scale surveys, such as the VACS and DHS; and (iii) selecting practices appropriate for young children. Guided by this, we implemented a survey experiment with four treatment arms that vary both how the data are collected (survey mode) and where the data are collected (survey location). Since English is the second language for most learners in Malawi, the survey was designed in three local languages: Chichewa, Tumbuka and Yao.

Survey mode. We used two survey administration modes: F2F and ACASI.[3] Under both modes, enumerators were trained to spend equal time building rapport with the child at the start

[3]See Figure A.3 in Appendix A for an illustration of how questions appear on the tablet when using ACASI.

of the interview to create a safe and comfortable environment. In the F2F mode, the enumerator asked the child directly about their experiences of school-related violence and recorded their responses on a tablet. Female enumerators interviewed girls, and male enumerators interviewed boys. In the ACASI mode, the enumerator provided the child with a tablet and headphones; the child listened to recorded questions and entered responses on a touchscreen. Before beginning, the enumerator explained how to use the device through short practice exercises, and remained nearby to answer questions, provide technical support, and monitor the child for signs of distress. To keep procedures comparable across survey modes, girls heard recordings in a female voice, while boys heard recordings in a male voice.

Survey location. We administered the survey either at the child's home or at their school, the two locations where most surveys are typically conducted. In both settings, the enumerators were trained to conduct the interviews in a space out of the listening range of others to ensure privacy.

Combining the two survey modes and two locations, children were randomly assigned to one of four treatment arms: F2F at home, F2F at school, ACASI at home, and ACASI at school. The control group consisted of children who completed the F2F survey at home, reflecting the status quo for violence surveys, including the DHS and VACS. These different combinations are expected to exogenously vary the degree of privacy provided during the survey, which is expected to affect respondents' barriers and fears associated with reporting experiences of violence. These barriers include fears of retaliation or re-victimization, concerns about social and relational consequences (such as judgment, stigma, or exclusion), and emotional inhibitors (such as shame, embarrassment, or guilt), all of which can affect a respondent's decision to disclose or conceal their experiences of violence when asked.

Relative to the F2F survey, ACASI removes direct interaction with the enumerator and reduces the perceived risks of being overheard, as the respondents directly self-report their experiences of violence on the tablet. This increase in privacy may help reduce fears related to retaliation, judgment, stigma, and exclusion, as well as emotional barriers such as shame and embarrassment, thereby affecting the likelihood of disclosure.

Survey location is also expected to influence reporting through its effects on perceived privacy and confidentiality. Differences between home and school settings may matter to the extent that the physical space, the physical distance to the perpetrator, and the people typically present in the environment (though not during the interview itself) shape children's perceptions of privacy

and confidentiality. For example, schools have generally larger spaces than homes and may offer more options for conducting the survey in spaces where privacy is salient, potentially reducing fears of retaliation and social consequences. Conversely, interviews conducted at home place respondents at greater physical distance from the perpetrators, as the survey exclusively focuses on violence perpetrated by peers or school staff. This greater distance may lower perceived risks of retaliation. Finally, the composition of people typically present in the environment may matter if children's internalized fears of judgment, exclusion, stigma, or embarrassment are stronger when family members or peers could potentially learn about their experiences of violence (Tourangeau and Yan 2007).

2.3 Data

2.3.1 Sample

The sample consisted of 6,009 students aged 8 to 12, distributed across 251 primary public schools in eight districts of Malawi: Blantyre, Chitipa, Dedza, Lilongwe, Mangochi, Mzimba, Ntcheu, and Zomba. The districts were selected randomly using proportionate-to-size sampling[4], based on the population of students enrolled in primary school. Within each district, the number of schools was determined according to the national distribution of schools, and mixed-sex public schools were selected randomly using proportionate-to-size sampling. Students within each school were selected randomly from school registers using the Ministry of Education Cohort Tracking Data, ensuring a balanced number of female and male students, as well as younger (8–9 years) and older students (10–12 years).[5]

We followed an individual-level randomization to allocate students to respond to the survey using (i) a F2F method at the student's home (control group), (ii) a F2F method at the student's school, (iii) an ACASI method at the student's home, and (iv) an ACASI method at the student's

[4]Malawi has 28 districts. Districts were randomly selected from a list of 19 districts in which the non-governmental organization YONECO operated. This was essential for the study, as YONECO was responsible for handling cases of violence reported in the survey that required referral services. Importantly, although our sample was restricted to districts where YONECO operated, comparisons of key statistics suggest that districts with and without YONECO presence are similar in observable characteristics (see Table A.1). A limitation of this analysis, however, is that we only have access to aggregated census measures and, given the low statistical power, we cannot rule out meaningful differences. Nevertheless, the overall pattern is reassuring.

[5]We originally aimed to sample 24 students in each school, resulting in 6,024 students. However, during the tracking phase of data collection—when enumerators confirmed the accuracy of age and gender data from the Cohort Tracking Data, and collected information on home addresses and caregivers' contact details—it was not possible to locate 15 children. As a result, in 21 schools the sample included between 20 and 23 students. In cases where the student was no longer enrolled in the school or the Cohort Tracking Data contained inaccurate information (for example, student age was outside the study's target range), enumerators were instructed to randomly select a replacement student from the hardcopy school registers, matching the same age, gender, and the first letter of the surname.

school. The randomization was stratified by school and student sex to ensure global balance within each stratum.[6] Assuming 80 percent power and a 10 percent significance level, our power calculations indicate that we can detect at least 3 to 5 percentage point differences in disclosure relative to the control group.[7]

We obtained data from 5,982 students. We could not administer the survey to 27 of 6,009 students (0.45 percent): 14 parents did not give consent, 5 children did not give assent, 7 children had a hearing or speaking disability, and 1 student could not be reached. The sample is balanced across our treatment arms, and we do not see evidence of differential attrition (see Table A.2 and A.3 in the Appendix).

2.3.2 Compliance

The survey software included features to monitor and minimize any deviations from the randomized treatment assignment. Regarding the survey mode, the software allowed us to prefill respondent data by assigned mode, ensuring that children assigned to the F2F arm were only listed in the F2F interface, and ACASI children only in the ACASI interface, thereby eliminating the risk that a child would be surveyed using the wrong method. Regarding survey location, we collected GPS coordinates during listing[8] and during the main data collection. We used this information to conduct daily checks to identify any violations of the protocol. We identified only one violation, in which the survey was conducted at the child's home instead of at the child's school. Aside from this case, we observed no further deviations from the assigned treatment.

2.3.3 Ethics and safeguarding

We received ethical clearance from the Institutional Review Board of the University of Malawi (Ref No: P.04/24/347). We implemented comprehensive safeguarding protocols to ensure the safety of all children during the study, minimize risks, and prevent any form of abuse by field staff. Enumerators and counsellors were thoroughly trained on study ethics and code of conduct, consent procedures and interview techniques, vicarious trauma and self-care, trauma-informed interviewing, rapport-building with children, and referral protocols and case man-

[6] The randomization was conducted using the Stata command `randtreat`, specifying the `misfits(global)` option. We adopted this approach because in some schools, the number of observations was not a multiple of the number of treatment arms. Following Carril (2017), we prioritized global balance of the misfits.

[7] Minimum detectable effects vary by outcome, ranging from 3 percentage points for sexual violence prevalence to 5 percentage points for corporal punishment and bullying. Power calculations were based on data on violence prevalence among 13- to 18-year-old children, due to the lack of data for younger children.

[8] Listing occurred one month before the survey to collect the GPS coordinates of children's homes.

agement[9]. Consent from the child's primary caregiver (obtained at the child's home) as well as the child's assent, was secured before administering the survey. All children were offered in-person counselling immediately after survey completion. Counsellors assessed each case, referring those requiring further attention to the non-governmental organization Youth-Net Counselling (YONECO), which specializes in child protection and operates the National Child Helpline Call Center. YONECO, in turn, involved local authorities, including social welfare offices and police victim units, when necessary. The details of our safeguarding procedures are discussed extensively in Evans, Lee, Smarrelli and Mhango (forthcoming).

2.4 Violence Outcomes

Our primary outcomes include emotional, physical, and sexual forms of violence perpetrated by students or members of the school staff. We describe each primary outcome below[10]:

Physical and Emotional bullying by peers: We measure physical and emotional patterns of repeated aggressive behaviour perpetrated by students during the month prior to the survey. Physical bullying is a dummy variable equal to 1 if the student experienced any of the following forms of aggression more than once in the past month: (i) being hit, kicked, or punched; (ii) being pinched or having ears or hair pulled; or (iii) having something thrown at them. Emotional bullying is a dummy variable equal to 1 if the student experienced any of the following more than once in the past month: (i) being made fun of or called names; (ii) being excluded from activities; (iii) having lies told or rumours spread about them; or (iv) being threatened.

Physical violence by the school staff: We construct a dummy variable equal to 1 if the student reported experiencing corporal punishment from any member of the school staff in the month prior to the survey. Corporal punishment includes spanking, beating, twisting ears, or any other form of hitting, whether with a hand or an object such as a stick.

Emotional violence by the school staff: We construct a dummy variable equal to 1 if the student reported that a member of the school staff humiliated, embarrassed, or threatened them in the month prior to the survey.

Sexual violence by peers or the school staff: We measure multiple forms of sexual violence, including both non-contact acts (e.g., comments of a sexual nature, nude pictures or

[9]See Figure A.4 in the Appendix for details on the referral protocol

[10]See Figure A.2 in Appendix A to review our survey questions.

videos) and contact acts (e.g., unwanted kissing or touching, forced sex). Using this information, we construct three dummy variables: (i) whether the child has ever experienced any form of sexual violence; (ii) whether they have experienced sexual violence involving contact; and (iii) whether they have experienced sexual violence without contact. For each outcome, we disaggregate by perpetrator type, distinguishing between students and school staff.

The survey questionnaire was adapted from validated instruments, including bullying items from international education assessment surveys such as PISA and TIMSS, and sexual violence items from the VACS and other studies focused on children (Barr et al. 2017). The questionnaire was piloted through focus group discussions and interviews with 100 children to assess comprehension. This included evaluating children's understanding of specific terms (e.g., "safety," "threaten," "embarrassment," "privacy"), the frequency scales, the time references (e.g., understanding of "month"), and the images used during the F2F and ACASI surveys to represent frequency and agreement scales.

In addition to our violence outcomes, we also pre-registered secondary outcomes to validate or assess the robustness checks of our results (Smarrelli 2024). The measures we used include a Social Desirability Scale, and an indicator to measure children's attention levels. We discuss these measures in Section 6.

3 Empirical Strategy

To estimate the differences in the disclosure of violence across treatment arms, we estimate:

$$y_{isd} = \beta_0 + \beta_1 T_{isd}^{FS} + \beta_2 T_{isd}^{AS} + \beta_3 T_{isd}^{AH} + \epsilon_{isd}, \tag{3.1}$$

where y_{isd} is the outcome of interest (self-reported violence) for the individual i in school s located in district d. T_{isd}^{FS} is a dummy variable that takes the value of 1 if the child was allocated to respond to the survey with the F2F mode at their school; T_{isd}^{AS} is a dummy variable that takes the value of 1 if the child was allocated to respond to the survey in the ACASI mode at their school; and, T_{isd}^{AH} is a dummy variable that takes the value of 1 if the child was allocated to respond to the survey in the ACASI mode at their home. For all binary outcomes, we use a logit model to estimate (3.1) and then calculate the corresponding average marginal effect associated with each treatment arm. We also adjust for multiple hypothesis testing (see Table

A.4 in the Appendix).

T_{isd}^{FS}, T_{isd}^{AS} and T_{isd}^{AH}, indicate if the survey mode and location of the survey matters for disclosure relative to F2F interviews conducted at the respondents' home. In addition to our main specification, to increase precision, we also run an additional specification controlling for pre-specified individual, family and school characteristics. Individual-level controls include sex and age dummies. Household-level controls include wealth index and parent's education, and school characteristics include school size.[11] We also estimate the impacts using Post Double Selection Lasso to select control variables and find that our results are not sensitive to the method used to select controls (see Appendix B).

Lower and Upper Bounds. All questions on experiences of violence included "don't know" and "decline" response options. Children assigned to respond to the survey using ACASI (whether at home or at school) were 3 percentage points more likely to select these options compared to those who responded using F2F. This occurred despite enumerators being trained—and regularly reminded—to read "don't know" and "decline" when administering the survey via F2F. This suggests that either some enumerators did not consistently read these response options or that, as Malloy and Stolzenberg (2019) suggest, children did not feel comfortable selecting them, partly due to power imbalances in adult–child interactions and the fact that children are not accustomed to exercising these options in a survey context. To account for "don't know" and "decline" responses, we estimate Manski–Horowitz lower and upper bounds. To construct the upper bound, we assign $y_{isd} = 1$ (child experienced violence) to all respondents in the treatment arms who answered "don't know" or "decline," and $y_{isd} = 0$ (child did not experience violence) to those in the control group. The lower bound is created under the opposite assumption. We find narrow bounds, indicating that the results are not sensitive to differences in "don't know" and "decline" responses (and hence to missing values in the outcome variables) between ACASI and F2F (see Section 4).

[11] We do not cluster standard errors. Abadie et al. (2023) outline two main reasons for clustering. First, clustering may be required for experimental design reasons when treatment assignment occurs at the cluster level. Second, clustering may be appropriate for sampling design reasons when the data are collected using a clustered sampling design intended to support inference about a broader population. Because treatment in our study is randomized at the individual level, we do not cluster standard errors.

4 Results

4.1 Physical and Emotional Violence

Tables 1 and 2 present the differences in the probability of disclosing experiences of physical and emotional violence from peers and the school staff, respectively, relative to the control group. Columns (1) and (5) show results without controls, while columns (2) and (6) include controls. Columns (3), (4), (7), and (8) include the Manski–Horowitz lower and upper bounds used to adjust for "don't know" and "decline" responses.

The results suggest that children are more likely to disclose experiences of physical and emotional violence when asked using ACASI, and that survey location does not explain differences in reporting these forms of violence.

For measures on bullying perpetrated by peers, we find that in the control group (F2F at home), 46 percent and 61 percent of children self-reported experiencing physical and emotional bullying, respectively. When we compare this group with the ACASI respondents at home or school groups, we observe statistically significant differences in the disclosure of peer-to-peer bullying. The disclosure of emotional bullying increased by 6 and 7 percentage points (corresponding to 10 and 11 percent increases), among those responding to the survey with ACASI at home and at school, respectively. Similar results are observed for physical bullying, for which disclosure increased by 10 and 9 percentage points (corresponding to 21 and 19 percent increases) when responding with ACASI at home and school, respectively.

Our design also allows us to explore whether survey location affects disclosure. We do this by comparing differences in reporting within each survey mode. We find that when administering the survey using F2F interviews at the child's school instead of the child's home, the disclosure of emotional and physical bullying increases by 4 and 5 percentage points, respectively. However, after adjusting for multiple hypothesis testing, this difference remains statistically significant only for measures related to emotional bullying (see Table A.4). Moreover, we note that among those responding to the survey with ACASI, survey location does not explain differences in bullying disclosure.

Turning to physical (corporal punishment) and emotional forms of violence perpetrated by school staff (e.g., teachers and headteachers), we find that location does not matter: disclosure rates are similar between the groups responding to the survey via F2F at home or school, as well as between those groups responding to the survey with ACASI at either location. However,

when comparing the control group to the ACASI group, we observe significant differences. In the ACASI group, regardless of the survey location, the disclosure of corporal punishment increased by 9 percentage points, while the disclosure of emotional violence increased by 20 percentage points (which corresponds to an 88 percent increase relative to the control group).

Relative to corporal punishment, the sharp rise in reports of emotional violence in the ACASI group may reflect differences in how these forms of violence are perceived, accepted and normalized. Although corporal punishment has been banned in many settings, including Malawi[12], it remains a common and socially accepted practice—one that parents endorse and children often internalise as a legitimate form of discipline (Lokot et al. 2020). Emotional violence, by contrast, is less openly discussed and acknowledged. As a result, children may feel more ashamed to disclose that a staff member spoke to them in a way that made them feel embarrassed, threatened, or humiliated. Another possibility is that some children misunderstood the question when responding via ACASI. However, enumerator reports and findings from focus groups conducted before the survey to test children's understanding of the terminology used in the questions suggest that it is unlikely that children misunderstood the question.

Table 1: Bullying by peers

	Emotional Bullying				Physical Bullying			
	(1)	(2)	(3) Lower Bound	(4) Upper Bound	(5)	(6)	(7) Lower Bound	(8) Upper Bound
ACASI Home	0.059***	0.059***	0.042**	0.069***	0.095***	0.095***	0.078***	0.107***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)
ACASI School	0.065***	0.065***	0.041**	0.079***	0.085***	0.085***	0.069***	0.099***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)
F2F School	0.038**	0.038**	0.036**	0.041**	0.048***	0.048***	0.045**	0.051***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)
Observations	5888	5888	5982	5982	5896	5896	5982	5982
Mean	0.608				0.456			
Pvalue: diff AH&AS	0.745				0.622			
Pvalue: diff AS&FS	0.124				0.042			
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Notes: The table presents estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group."Pvalue: diff AH&AS" refers to differences between the ACASI group that responded to the survey at school or home. "Pvalue: diff AS&FS" refers to differences between those who responded to the survey in school, using F2F or ACASI. Columns (2) and (6) include all individual, family and school pre-specified covariates. Columns (3), (4), (7) and (8) show the Manski–Horowitz lower and upper bounds used to adjust for "don't know" and "decline" responses. Robust standard errors are shown in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[12]Corporal punishment is unlawful in schools under article 19 of the Constitution. Yet, the Education Act does not explicitly prohibit the use of corporal punishment in schools. Only the National Education Standards and the Teachers' Code of Conduct guidelines discourage teachers from using corporal punishment in schools.

Table 2: Emotional & Corporal Punishment by Staff

	Emotional				Corporal Punishment			
	(1)	(2)	(3) Lower Bound	(4) Upper Bound	(5)	(6)	(7) Lower Bound	(8) Upper Bound
ACASI Home	0.206***	0.206***	0.191***	0.223***	0.088***	0.088***	0.072***	0.101***
	(0.017)	(0.017)	(0.017)	(0.016)	(0.018)	(0.018)	(0.018)	(0.018)
ACASI School	0.207***	0.208***	0.190***	0.228***	0.093***	0.093***	0.070***	0.111***
	(0.017)	(0.017)	(0.017)	(0.016)	(0.018)	(0.018)	(0.018)	(0.018)
F2F School	0.023	0.023	0.018	0.028	-0.009	-0.009	-0.014	-0.003
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Observations	5873	5873	5982	5982	5876	5876	5982	5982
Mean	0.233				0.473			
Pvalue: diff AH&AS	0.903				0.775			
Pvalue: diff AS&FS	0.000				0.000			
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Notes: The table presents estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group."Pvalue: diff AH&AS" refers to differences between the ACASI group that responded to the survey at school or home. "Pvalue: diff AS&FS" refers to differences between those who responded to the survey in school, using F2F or ACASI. Columns (2) and (6) include all individual, family and school pre-specified covariates. Columns (3), (4), (7) and (8) show the Manski–Horowitz lower and upper bounds used to adjust for "don't know" and "decline" responses. Robust standard errors are shown in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.2 Sexual Violence

Figure 1 shows the percentage of children who reported experiencing different forms of sexual violence in the control group, from sexual comments about the child's body to forcing, persuading, or enticing a child to take part in sexual activities, which may or may not involve contact. We asked about these forms of violence in the order presented in Figure 1, differentiating between whether the perpetrator was a peer or an adult (member of the school staff).

Figure 1: Percentage of children who self-report being victims of sexual violence

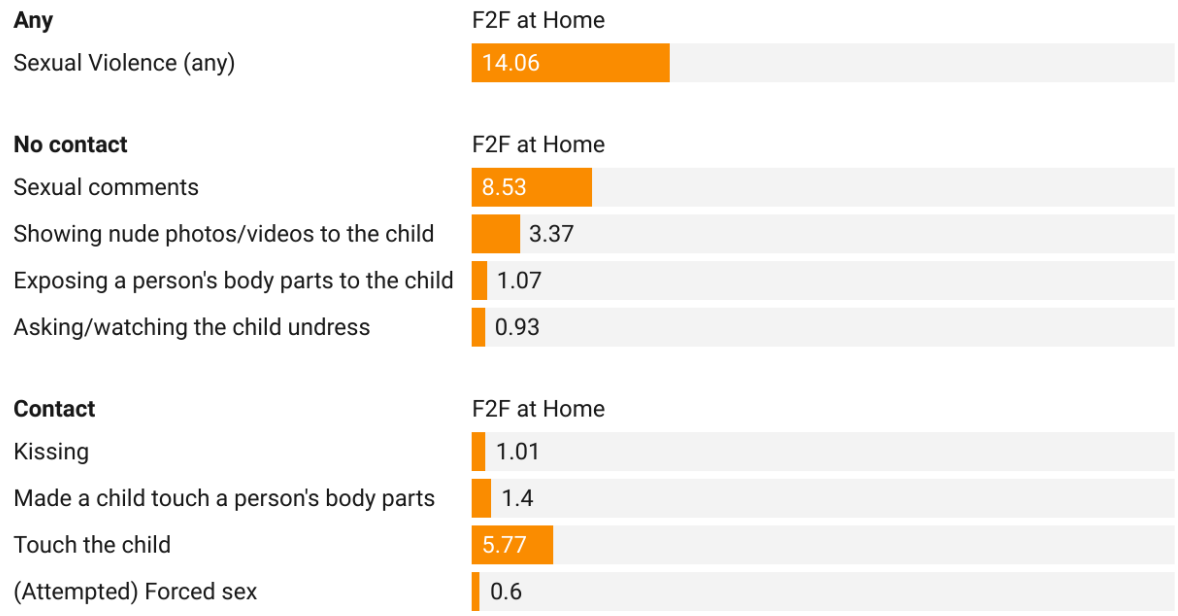

Table 3 shows that, relative to the control group, the disclosure of sexual violence increased by 13 and 17 percentage points when the survey was administered via ACASI at home and school, respectively. This corresponds to a 91–118 percent difference relative to the control group, driven by differences in the violence perpetrated by adults, including members of the school staff or other adults on the way to or from school. This finding would suggest that the higher privacy afforded by ACASI is especially important when asking about abuse from adult perpetrators. Results are similar when disaggregating the data by whether the form of sexual violence involved contact or not (see Table A.5 and A.6 in Appendix A).

Table 3: Sexual Violence

	Sexual Violence			By Staff			By Peers		
	(1)	(2) Lower Bound	(3) Upper Bound	(4)	(5) Lower Bound	(6) Upper Bound	(7)	(8) Lower Bound	(9) Upper Bound
ACASI Home	0.128***	0.125***	0.137***	0.152***	0.146***	0.169***	0.076***	0.073***	0.090***
	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.015)
ACASI School	0.167***	0.163***	0.179***	0.173***	0.167***	0.193***	0.102***	0.096***	0.121***
	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)
F2F School	0.078***	0.078***	0.080***	0.044***	0.040**	0.047***	0.067***	0.067***	0.071***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.016)	(0.018)	(0.015)	(0.015)	(0.016)
Observations	5942	5982	5982	5925	5982	5982	5922	5982	5982
Mean	0.141	0.141	0.141	0.038	0.039	0.037	0.124	0.124	0.124
Pvalue: diff AH&AS	0.006	0.009	0.003	0.019	0.024	0.010	0.063	0.086	0.023
Pvalue: diff AS&FS	0.000	0.000	0.000	0.000	0.000	0.000	0.010	0.026	0.000
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group."Pvalue: diff AH&AS" refers to differences between the ACASI group that responded to the survey at school or home. "Pvalue: diff AS&FS" refers to differences between those who responded to the survey in school, using F2F or ACASI. Columns (1), (4), and (7) include all individual, family and school pre-specified covariates. Columns (2), (3), (5), (6), (8) and (9) show the Manski–Horowitz lower and upper bounds used to adjust for "don't know" and "decline" responses. Robust standard errors are shown in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We also find that the location of the survey affects disclosure. Among those surveyed using F2F, administering the survey in school increased the disclosure of sexual violence by 8 percentage points, corresponding to a 55 percent increase in reporting. Among those surveyed using ACASI, the disclosure of violence was also higher in the school setting. The higher disclosure of sexual violence in the school relative to the home setting may reflect differences in the respondents' perceived levels of privacy and anonymity. As explained in Section 2.2, three interconnected factors might affect these perceived levels of privacy when answering survey questions on violence: (i) the physical space, (ii) the physical distance to the perpetrator, and (iii) the people typically present in the environment (though not during the interview itself).

All the questions in our survey asked about school-related violence, focusing on incidents perpetrated by peers or the school staff. In this context, if the physical distance to the perpetrator mattered the most, we would have expected to observe higher disclosure of violence

among the groups allocated to respond to the survey at home. Given that we do not observe this, it is likely that the other two factors matter more in our context.

Our survey included a set of questions asking the enumerator to describe the physical space where the survey was conducted. Although enumerators in both school and home settings were instructed to ensure surveys were conducted out of the listening range of others, the characteristics of the physical spaces varied. We find that, at home, 4 percent of surveys were conducted inside a room (primary indoor), 87 percent in common passage areas or entry zones (specifically, at the front door or veranda), and 9 percent in outdoor spaces (e.g., the house yard). In schools, 21 percent of surveys were conducted inside a room (mainly classrooms), 30 percent in common passage areas or entry zones (e.g., in a school hallway or near the main entrance), and 49 percent in outdoor spaces (e.g., the school yard). There was more variation in the use of school spaces, suggesting that enumerators could use different options to ensure the survey was administered out of hearing range of others. In contrast, children's homes had, on average, only two rooms, and the fact that the house's front door was the most common place used to administer the interview suggests that this space was likely the most private setting available for ensuring that the interview was conducted out of the hearing range of others.

In Table 4, we explore the association between the physical spaces used to conduct the survey and the violence outcomes. We include in the regression two dummy variables that capture whether the survey was conducted inside a room or in a common passage area or entry zone, and we use as a reference category whether the interview was conducted in an outdoor space further away from the physical structure of the house or school (e.g., in the school or home yard). We estimate this specification separately for the children allocated to respond to the survey in their homes and schools. We find that for physical or emotional forms of violence perpetrated by peers or school staff, the physical space is not correlated with violence disclosure. For outcomes related to sexual violence, the results suggest that, in the home, conducting interviews inside a room or in common passage areas is associated with lower disclosure of sexual violence relative to administering the survey in outdoor spaces. By contrast, in the school, we do not observe statistically significant differences.

The results suggest that in schools, it is possible to use a variety of spaces that allow enumerators to signal more clearly that privacy is secured (as enumerators may be able to find spaces that are more clearly outside the listening range of others), relative to the home, where space is more constrained. In fact, qualitative reports from the enumerators and counselors on

Table 4: Physical Spaces within Home and School

	(1) Home	(2) School
Panel A: Sexual Violence		
Interview conducted: Inside a room	-0.079*	0.033
	(0.045)	(0.020)
Interview conducted: Common Passage Area/Entry Zone	-0.041*	-0.000
	(0.024)	(0.019)
N	2967	2962
Panel B: Physical or Emotional Bullying from Peers		
Interview conducted: Inside a room	-0.050	0.011
	(0.052)	(0.022)
Interview conducted: Common Passage Area/Entry Zone	-0.050	-0.002
	(0.031)	(0.019)
N	2915	2917
Panel C: Physical or Emotional Violence from School Staff		
Interview conducted: Inside a room	0.016	0.015
	(0.055)	(0.024)
Interview conducted: Common Passage Area/Entry Zone	-0.005	0.030
	(0.032)	(0.021)
N	2910	2897

Notes: The table presents three panel of results showing the estimated coefficients of physical spaces within home and school on the disclosure of violence. Column (1) shows the estimated effects for those interviewed at home, while column (2) shows the estimated effects for those interviewed at school. All regressions include a control for survey mode. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the data collection team also indicated that children appeared to be less stressed or more relaxed in the school setting than at home. This may indicate that children may fear or worry more about being overheard at home, potentially due to perceived social consequences in the home or community setting, including judgment, disbelief, gossip, or disruption of family relationships (Nurjannah et al. 2024). Taken together, our analysis would support the hypothesis that, in our context, the physical spaces used to administer the survey and, potentially, concerns that family members might overhear or become aware of children's experiences of violence, play a greater role in explaining our results than other channels, such as the physical distance from the perpetrator.

5 Characteristics of Misreporting

In this section, we explore the type of person who might be more likely to misreport their experiences of violence. Following Cullen (2023) approach, we run the following specification for each of our treatment arms separately:

$$Y_i = \gamma_0 + \gamma_1 X_i + \epsilon_i, \tag{5.1}$$

where Y_i refers to the violence outcome and X_i refers to a vector of individual, family, and context-level characteristics that are likely to influence misreporting or that constitute violence risk factors. We then compare the estimated coefficient across measurement methods. If the coefficients are statistically significantly different, this would indicate that the characteristic is sensitive to the measurement method, providing suggestive evidence of systematic misreporting.[13]

5.1 Results

In Table 5, we show the relationship between child, family, and context-level characteristics and the likelihood of emotional, physical and sexual violence by survey method. We focus on this presentation of results rather than showing results for the four treatment arms, as we do not observe significant differences by survey location (see Tables A.10 to A.15 in Appendix A).

The results suggest that the analysed characteristics are correlated with the likelihood of emotional violence in similar ways across the survey modes, whereas we observe important differences for physical and sexual violence.

Consistent with existing evidence (Scheithauer et al. 2006; López-Castro et al. 2023; Zhou et al. 2023), children's age is negatively correlated with the likelihood of physical violence, suggesting that children aged 10 to 12 are less likely to experience physical violence relative to children aged 8 and 9. We observe, however, that the ACASI and F2F estimated coefficients are different. Specifically, when asked using F2F, older children appear to be more likely to underreport their experience of physical violence, which can be interpreted as a sign of potential systematic misreporting. This pattern is consistent with the idea that, as children grow older, they internalize norms related to embarrassment and shame, making them less willing to disclose such experiences in direct, face-to-face questioning. Interestingly, we do not observe similar age-related differences across survey methods for emotional or sexual violence, suggesting that the role of internalized norms on reporting and its consequences vary across forms of violence.

[13]This approach is similar to estimating equation 3.1 with interaction terms between the treatment variable and the covariate of interest. Tables A.7, A.8, and A.9 report results using this approach, which yield similar findings.

We also find that the coefficient for gender between the ACASI and F2F arms is different. Boys assigned to the ACASI treatment arms are more likely to report sexual violence than those assigned to the F2F arms. This would suggest a higher likelihood of systematic misreporting among boys, and shows that the higher privacy provided by ACASI is particularly relevant to reduce misreporting for them. This would be consistent with evidence showing that boys are less likely to speak about their experiences of violence relative to girls (Hohendorff et al. 2017; Hietamäki et al. 2024). Notably, we observe method-related differences by gender only for sexual violence. It is possible that the feelings of embarrassment, shame and disempowerment for being a victim of sexual abuse, as well as fears of not being believed or taken seriously due to harmful myths supporting that men are not sexually abused, or restrictive norms around traditional masculinity, become more pronounced barriers to reporting sexual abuse relative to emotional and physical abuse (Feder et al. 2010; Turchik and Edwards 2012).

The analysis in this section highlights that the data collection method matters to minimize systematic misreporting. Failing to assess which method is most appropriate for each context and population subgroups increases the likelihood of obtaining biased estimates when analysing the impact of violence-related interventions.

6 Discussion and Robustness

The results indicate that survey mode matters, with higher disclosure of violence with ACASI relative to F2F. Survey location also explains differences in reporting, particularly for sexual violence. In this section, we further examine the reliability and robustness of these findings. We do this by examining social desirability bias, enumerator effects, attention lapses, straightlining, and question misunderstandings.

6.1 Social Desirability Bias

ACASI provides greater privacy, as children listen to a recording and select the image that best represents their response. The enumerator remains nearby to address questions, but does not directly ask the questions, adding an extra layer of confidentiality and likely reducing fears of breaches of privacy. To further assess whether this additional privacy improves the accuracy of reporting, we examine differences in social desirability bias (SDB) between ACASI and F2F.

SDB refers to the tendency to underreport stigmatized or socially undesirable behaviors

Table 5: Correlates of Misreporting

	Emotional Violence		Diff (p-value)	Physical Violence		Diff (p-value)	Sexual Violence		Diff (p-value)
	ACASI	F2F		ACASI	F2F		ACASI	F2F	
	(1)	(2)	(1) - (2)	(1)	(2)	(1) - (2)	(1)	(2)	(1) - (2)
Individual Characteristics									
Male student	0.035**	0.056***	0.382	0.024	0.020	0.894	0.095***	0.024*	0.001
	(0.016)	(0.017)		(0.017)	(0.017)		(0.016)	(0.014)	
Old cohort [10 to 12]	-0.019	-0.029*	0.674	-0.014	-0.061***	0.050	0.058***	0.060***	0.934
	(0.016)	(0.017)		(0.017)	(0.017)		(0.016)	(0.014)	
Family Characteristics									
Wealth Index	-0.037	-0.073	0.594	-0.047	0.044	0.180	0.060	-0.020	0.186
	(0.047)	(0.049)		(0.047)	(0.049)		(0.047)	(0.038)	
Caregiver: complete secondary or more	0.005	-0.011	0.664	0.030	-0.062**	0.016	-0.014	-0.042**	0.389
	(0.026)	(0.028)		(0.026)	(0.028)		(0.026)	(0.021)	
Main practised religion: Christian	0.057**	0.042*	0.263	0.070***	0.041*	0.403	0.045**	0.029	0.578
	(0.023)	(0.024)		(0.024)	(0.024)		(0.022)	(0.019)	
Environment: Characteristics									
High Density School	0.027**	0.006	0.246	-0.020	-0.015	0.766	-0.008	-0.007	0.943
	(0.013)	(0.013)		(0.013)	(0.013)		(0.013)	(0.010)	
High Density City	-0.016	-0.046**	0.252	-0.054***	-0.040**	0.607	0.016	0.029*	0.578
	(0.018)	(0.019)		(0.018)	(0.019)		(0.018)	(0.016)	

Notes: The Table shows the relationship between child, family, and context-level characteristics and the likelihood of emotional, physical, and sexual violence by survey method. The columns titled 'Diff (pvalue)' show the two-sided p-value test difference between the ACASI and F2F groups. p-values below 0.10 indicate that the characteristic is sensitive to the measurement method, providing suggestive evidence of systematic misreporting. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

and to overreport socially approved ones. When applied to violence, the stigma associated with being a victim, along with social norms that encourage victim-blaming, can bias reporting downward, especially among respondents prone to provide socially acceptable answers. We measure this using the Children's Social Desirability Short (CSD-S) scale, originally developed by Crowne and Marlowe (1960) and adapted for children by Miller et al. (2016). The scale includes 14 yes/no items, such as "Have you ever felt like saying unkind things to a person?", "Do you ever say anything that makes somebody else feel bad?", and "Do you always listen to your parents?". We compute the total number of socially desirable responses and construct a binary indicator equal to 1 if the score is above the median.

Table 6, Columns 1 and 2, shows a lower socially desirability bias among those who responded to the survey using ACASI, providing supporting evidence that higher privacy reduces the likelihood of providing socially desirable answers, and hence that systematic misreporting is less likely in the ACASI arm.

6.2 Enumerator Effects

A related concern is whether results are driven by particular enumerators. If certain enumerators systematically elicited more or fewer reports of violence, due to differences in rapport-building, tone, or interviewing style, this could confound the comparison between ACASI and F2F. We assess this by re-estimating our main specifications with enumerator fixed effects. The results remain consistent across all outcomes (see Tables A.20 and A.21 in the Appendix), suggesting that enumerator-level variation does not drive our findings.

6.3 Lack of Attention

To assess potential measurement error from fatigue or low motivation, we embedded two attention-check questions in the survey: (i) "My teacher is older than me" (asked at the start), and (ii) "Is 1+1=3?" (asked at the end). Each has a unique correct response and does not capture cognitive ability, making them standard checks of attentiveness. Respondents fail the attention check if they answer "totally disagree" or "disagree" to the first item, and "Yes" to the second. In the control group, 9 percent of children failed at least one attention check question. Table 6, Columns 3 and 4, shows that lapses in attention are 3 to 5 percentage points more likely among respondents assigned to ACASI. This might be a concern, particularly if we believe these are non-random errors.

We examine whether this affects our results in two ways. First, we compare children who failed the attention checks to those who did not, and note that inattentive children are mainly young boys with a lower wealth index. We remove these children from the sample and observe that the results remain unchanged, suggesting that inattentive respondents do not drive our findings (see Table A.17 in the Appendix). Second, we include the variable for failing the attention checks as a covariate. While this constitutes a bad control, the fact that results remain consistent provides suggestive evidence that failing the attention checks does not necessarily imply that responses on violence are less reliable (see Table A.18 in Appendix A).

Table 6: Social Desirability Bias and Attention Checks

	SDB		Fail Attention Check	
	(1)	(2)	(3)	(4)
ACASI Home	-0.076***	-0.076***	0.032***	0.032***
	(0.017)	(0.017)	(0.012)	(0.012)
ACASI School	-0.070***	-0.070***	0.046***	0.047***
	(0.017)	(0.017)	(0.011)	(0.011)
F2F School	-0.006	-0.006	-0.007	-0.006
	(0.016)	(0.016)	(0.013)	(0.012)
Observations	5982	5982	5982	5982
Mean	0.352		0.092	
Controls	No	Yes	No	Yes

Notes: The table presents estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group. Columns (2) and (4) include all individual, family and school pre-specified covariates. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.4 Straightlining

Straightlining refers to a situation in which a respondent provides the same answer for a series of questions that form part of the same grid or survey module. Evidence of straightlining raises concerns about data quality. To assess this, we focus on six survey modules in which questions share identical response options and are asked consecutively. For each module, we calculate the variability of responses and create a dummy variable equal to one if the standard deviation is equal to zero. We aggregate the dummy variables to calculate the number of modules with suggestive evidence of straightlining and find that the median and mean of this variable are 1 and 1.6, respectively. Based on this, we construct a dummy variable that takes the value 1 if straightlining occurs in at least one module and 0 otherwise. As shown in Table A.19, Column 1, the likelihood of straightlining is low and, if anything, negative relative to the control group. Using the number of modules with evidence of straightlining as the outcome (Column 2), we

find that it is slightly higher in the ACASI arm, although the magnitude of the coefficient is small.

Our measure of straightlining is, however, imperfect. A key limitation is that we may incorrectly assume that respondents should not legitimately give the same answer to all items. This is particularly relevant in our context, as all but one of the modules asked about violence, and providing the same response across items may still be valid. For example, in the sexual violence module, we observe a higher number of students responding "no" to all items in the F2F arm compared with the ACASI arm. Similarly, within the ACASI arm, more students reported they experience bullying "lots of times" across all the items of the question relative to the F2F arm. Both can be valid responses, and therefore our straightlining measure must be interpreted with caution.

Importantly, neither our straightlining dummy nor the count of modules with evidence of straightlining is significantly correlated with survey duration. The difference in survey duration between respondents with and without straightlining is approximately three minutes in both the ACASI and F2F arms. We therefore do not view straightlining as a substantive threat to data quality in this study, as we would expect surveys to be completed faster if students intentionally decided to give the same response within survey modules.

6.5 Potential Misunderstanding of Questions

Prevalence and nature. We piloted the survey with 100 children and organised two focus groups—one with children aged 8 to 9 and one with children aged 10 to 12— to assess comprehension of key concepts and response options. These steps allowed us to make revisions to the survey instrument before data collection. Nevertheless, some children may still have misunderstood some questions during the survey.

Identifying which respondents misunderstood survey questions is inherently difficult. We are able to examine this issue within the subsample of children referred for counselling, using anonymized counsellor reports indicating whether a case required additional follow-up and whether the counsellor believed the child had misunderstood a survey question. Although this subsample differs from the full sample on observable characteristics, it is a relevant subsample to explore suggestive evidence on the prevalence and nature of misunderstandings (see Table A.22).

All children were offered counselling at the end of the survey, and those requesting counselling

met with one after the survey (see Figure A.4 in the Appendix). In addition, among those who reported experiencing forced sex, even if the child declined counselling, our safeguarding protocol required a counsellor to assess whether a referral to the appropriate authorities was necessary. Consequently, all children who reported forced sex met with a counsellor.

Overall, 1,566 out of 5,982 surveyed children received counselling. Approximately 10 percent of these cases involved a potential misunderstanding of at least one survey question. Most misunderstandings concerned the forced-sex question.[14] In most such cases, the child appeared either to have interpreted forced touching as forced sex or to have misunderstood the question more broadly, with the counsellor concluding that no form of sexual violence had occurred.

Distribution of misunderstandings by treatment arms and student characteristics. A potential concern is that the higher disclosure observed under ACASI may reflect false positives arising from misunderstandings rather than genuine reductions in underreporting. We assess this possibility in two steps. First, we examine whether children who misunderstood the questions differ systematically from those who did not, using child and household characteristics. We find that children who misunderstood questions do not differ significantly from those who did not, with one exception: boys were more likely to misunderstand the questions (see Table A.23).

Second, we investigate whether the likelihood of misunderstanding a question differs by treatment assignment. Because the analysis relies on the selected counselling subsample, the results should be interpreted as descriptive rather than causal. We find that assignment to the ACASI arm is associated with an approximately 6 percentage point increase in the likelihood of misunderstanding a question (Table A.25, Column 1). Although children in the ACASI arm were able to seek clarification from enumerators if they had questions or doubts about a question, misunderstandings may have been easier to identify and resolve directly during F2F interviews.

We explore whether this difference could reflect compositional differences across treatment arms within the counselling subsample and find this is not the case. Although the counselling subsample is non-randomly selected, it remains balanced across treatment arms on individual and household characteristics (see Table A.24). We then assess whether misunderstandings are concentrated among particular observable subgroups by estimating regressions that interact

[14]For all cases reporting forced sex in the full sample, and for 65 percent of cases reporting other forms of sexual violence, we were able to assess whether the child had misunderstood the question.

ACASI assignment with key respondent characteristics, including sex, age, caregiver education, and household wealth. None of the interaction terms is statistically significant (Table A.25, Columns 2–6). This provides suggestive evidence that the higher rate of misunderstandings observed under ACASI is unlikely to be driven by systematic differences across observable characteristics. However, these results should be interpreted cautiously. In the interacted specifications, the coefficient on ACASI represents the effect for the omitted or reference group rather than the overall average effect, and it becomes imprecisely estimated. Moreover, the small and selected nature of the counselling subsample limits statistical power, so the results from this analysis should not be over-interpreted.

Most importantly, our main results remain largely unchanged after accounting for potential misunderstandings (Table 7)[15]. Specifically, the results remain unchanged when comparing the control group to F2F at school, and the estimated ACASI coefficients decline by approximately 2 percentage points after adjusting for potential misunderstandings. The main overall finding remains the same: disclosure of sexual violence under ACASI is 78 to 100 percent higher than under F2F, even after accounting for potential misunderstandings. These findings suggest that, while misunderstandings may contribute some false positives, they are unlikely to explain the higher disclosure observed under ACASI.

[15] Analysis of the correlates of misreporting also holds when adjusting for potential misunderstandings; see Table A.26

Table 7: Sexual Violence Adjusted

	All		Staff		Peer	
	(1)	(2) Adjusted	(3)	(4) Adjusted	(5)	(6) Adjusted
ACASI Home	0.128***	0.102***	0.151***	0.122***	0.076***	0.054***
	(0.016)	(0.016)	(0.015)	(0.014)	(0.015)	(0.014)
ACASI School	0.167***	0.138***	0.173***	0.139***	0.101***	0.085***
	(0.016)	(0.015)	(0.015)	(0.013)	(0.015)	(0.014)
F2F School	0.078***	0.071***	0.043***	0.034**	0.067***	0.061***
	(0.017)	(0.016)	(0.017)	(0.015)	(0.015)	(0.014)
Observations	5942	5942	5925	5925	5922	5922
Mean	0.141	0.130	0.038	0.036	0.124	0.114
Pvalue: diff AH&AS	0.006	0.008	0.019	0.056	0.063	0.015
Pvalue: diff AS&FS	0.000	0.000	0.000	0.000	0.010	0.058
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group."Pvalue: diff AH&AS" refers to differences between the ACASI group that responded to the survey at school or home. "Pvalue: diff AS&FS" refers to differences between those who responded to the survey in school, using F2F or ACASI. All columns include all individual, family and school pre-specified covariates. Columns (2), (4) and (6) include an outcome variable that accounts for any cases where the question on sexual abuse was misunderstood. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.6 Summary

In the section above, we explored the reliability and robustness of our results and examined potential challenges related to both over- and underreporting. In the absence of an external ground truth, it is difficult to determine whether differences in disclosure reflect reductions in underreporting or increases in overreporting. Although both sources of measurement error are possible, research on violence suggests that higher disclosure is more likely to reflect the truth than false positives, as the main risk is generally underreporting (Tourangeau and Yan 2007; Peterman et al. 2025). This is especially likely to be the case in surveys with young children, for whom reporting errors are more likely to arise from misunderstandings than incentives to over-report. Our robustness analysis provides reassurance on this point.

Taken together, the findings suggest that ACASI reduces underreporting relative to F2F interviewing, consistent with evidence that ACASI elicits fewer socially desirable responses, which are more likely to generate systematic measurement error. At the same time, we find some evidence of measurement error in ACASI, including higher rates of attention lapses and question misunderstandings. However, these errors appear more likely to reflect random rather than systematic measurement error, and our estimates remain similar after accounting for them. Overall, this suggests that the higher disclosure observed under ACASI is unlikely to be driven

primarily by false positives.

7 Cost-Effectiveness

We combine data on estimated treatment effects and program costs to assess the cost-effectiveness of each data collection method. In the measurement context, effectiveness is defined as the ability of a method to provide reliable measures of violence, that is, with the lowest measurement error. We interpret this as the method yielding the highest disclosure of violence relative to the control group, assuming that high disclosure reflects accurate reporting and reductions in misreporting, mainly stemming from underreporting. This aligns with research on sensitive topics, which suggests that respondents generally underreport socially undesirable situations and behaviors (such as violence) and overreport socially desirable ones (such as seat belt use) (Tourangeau and Yan 2007).

To operationalize cost-effectiveness, we focus on one outcome—sexual violence—given its sensitivity and policy relevance. The cost structure varies across treatment arms. Relative to F2F surveys, the main additional costs of ACASI arise from software for recording the audio files, the salaries of staff responsible for recording and editing, and the purchase of equipment such as headphones. By contrast, location affects costs in the opposite direction: school-based implementation reduces transport costs, since all children are surveyed in the same place. Our cost data indicate that ACASI was USD 1 more expensive per child than F2F interviews, while school-based surveys were USD 3.99 cheaper per child than home-based surveys.

Using these data, we compute the cost-effectiveness ratio by dividing the treatment effect by the costs of each arm. The results show that increasing the probability of reporting sexual violence by one percentage point costs USD 3.67 when collecting data using ACASI at school, USD 5.10 with ACASI at home, and USD 7.69 with F2F at school. These findings suggest that ACASI is not only effective in eliciting disclosure of violence but also cost-effective[16].

8 Conclusion

Reducing misreporting of sensitive issues is crucial for generating reliable statistics and improving our understanding of what works to prevent violence against children. Existing evidence on

[16]The estimates would remain similar if we use the measure of sexual violence adjusted for potential misunderstandings. See Appendix C, Table A.31

tools that mitigate misreporting has largely focused on adults, so little is known regarding how these approaches perform among younger populations. We address this gap by implementing a large survey experiment with around 6,000 children aged 8 to 12 in Malawi.

Our findings demonstrate that survey mode and location affect the disclosure of violence against children, with lack of privacy being a key driver of misreporting. ACASI significantly increases disclosure of physical, emotional, and sexual violence relative to F2F interviews, with the largest differences observed for sexual violence. School-based administration further enhances reporting of sexual violence, with quantitative and qualitative evidence suggesting that the physical spaces used to administer the survey, as well as concerns about family members overhearing or becoming aware of children's experiences of violence, partly explain this finding. Our robustness analysis shows that ACASI leads to lower socially desirable responses, and that despite a higher likelihood of misunderstanding certain questions or failing attention checks in ACASI, results are not driven by these issues. We also find that, despite slightly higher per-child costs, ACASI is cost-effective relative to F2F surveys.

These findings have implications for researchers and policymakers collecting data on school-related violence. They provide evidence on survey methods that can be used to collect information on sensitive topics from children aged 8 to 12. More broadly, they highlight the importance of carefully assessing how violence data are collected in large-scale national and international surveys, as well as in research studies. We cannot directly extrapolate our results to assess the magnitude of potential misreporting in other surveys. Yet, if our results did extend to older populations, this would suggest that existing surveys that rely on face-to-face interviewing may be underestimating the prevalence of violence. As an illustrative example, we apply our estimated misreporting rates to data from the VACS in Malawi, administered through face-to-face interviews among children aged 13 to 17. We focus on two measures: (i) forced or pressured sex, and (ii) forced, pressured, or attempted sex perpetrated by peers or teachers. This exercise implies that estimated prevalence could be between 2 and 7 times higher than reported, depending on the outcome definition[17]. This result should be read as suggestive only, given differences in age composition and question framing. Nevertheless, it illustrates the importance of further examining which survey methods most effectively enhance privacy and reduce potential misreporting across different contexts and age groups.

Our results alone are insufficient to establish ACASI as a gold standard for measuring

[17]See Appendix D for more information on our assumptions.

violence against young children. Rather, they show that—relative to face-to-face interviewing—ACASI helps reduce misreporting. Future research in other contexts, comparing ACASI with alternative survey methods, as well as exploring the role of survey location in violence disclosure using both quantitative and qualitative approaches, would be valuable for validating and extending our findings.

References

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge (2023). When Should You Adjust Standard Errors for Clustering? *The Quarterly Journal of Economics 138*(1), 1–35. _eprint: https://academic.oup.com/qje/article-pdf/138/1/1/47915437/qjac038.pdf.

Agüero, J. M. and V. Frisancho (2022). Measuring Violence against Women with Experimental Methods. *Economic Development and Cultural Change 70*(4), 1565–1590. Publisher: The University of Chicago Press.

Barr, A. L., L. Knight, I. Frana-Junior, E. Allen, D. Naker, and K. M. Devries (2017). Methods to increase reporting of childhood sexual abuse in surveys: the sensitivity and specificity of face-to-face interviews versus a sealed envelope method in Ugandan primary school children. *BMC International Health and Human Rights 17*, 4.

Belloni, A., V. Chernozhukov, and C. Hansen (2014). High-Dimensional Methods and Inference on Structural and Treatment Effects. *Journal of Economic Perspectives 28*(2), 29–50.

Blair, G., A. Coppock, and M. Moor (2020). When to worry about sensitivity bias: A social reference theory and evidence from 30 years of list experiments. *American Political Science Review 114*(4), 1297–1315.

Blattman, C., J. Jamison, T. Koroknay-Palicz, K. Rodrigues, and M. Sheridan (2016). Measuring the measurement error: A method to qualitatively validate survey data. *Journal of Development Economics 120*, 99–112.

Boudreau, L. E., S. Chassang, A. Gonzalez-Torres, and R. M. Heath (2023). Monitoring Harassment in Organizations. *National Bureau of Economic Research Working Paper 3101*.

Bulte, E. and R. Lensink (2019). Women's empowerment and domestic abuse: Experimental

evidence from Vietnam. *European Economic Review 115*, 172–191.

Carril, A. (2017). Dealing with misfits in random treatment assignment. *Stata Journal 17*(3), 652–667. Publisher: StataCorp LLC.

Cullen, C. (2023). Method Matters: The Underreporting of Intimate Partner Violence. *The World Bank Economic Review 37*(1), 49–73.

De Weerdt, J., J. Gibson, and K. Beegle (2020). What can we learn from experimenting with survey methods? *Annual Review of Resource Economics 12*(1), 431–447.

Evans, D. K., S. Hares, G. Smarrelli, and D. Wu (2025). When the data you have aren't the data you need: The availability of school-related violence data in low- and middle-income countries. *World Development 188*, 106919.

Feder, J., R. F. Levant, and J. Dean (2010). Boys and violence: A gender-informed analysis. *Psychology of Violence 1*(S), 3–12. Place: US Publisher: Educational Publishing Foundation.

Gibson, J., K. Beegle, J. De Weerdt, and J. Friedman (2015). What does variation in survey design reveal about the nature of measurement errors in household consumption? *Oxford Bulletin of Economics and Statistics 77*(3), 466–474.

Gilligan, D. O., M. Hidrobo, J. Leight, and H. Tambet (2025). Using a list experiment to measure intimate partner violence: cautionary evidence from Ethiopia. *Applied Economics Letters 32*(11), 1594–1600. Publisher: Routledge _eprint: https://doi.org/10.1080/13504851.2024.2308579.

Hietamäki, J., M. Husso, T. Arponen, and H.-M. Lahtinen (2024). Differences Between Girls and Boys in the Disclosure of Sexual Violence. *Journal of Interpersonal Violence 39*(11-12), 2629–2654.

Hillis, S., J. Mercy, A. Amobi, and H. Kress (2016). Global Prevalence of Past-year Violence Against Children: A Systematic Review and Minimum Estimates. *Pediatrics 137*(3). Publisher: American Academy of Pediatrics.

Hohendorff, J. V., L. F. Habigzang, and S. H. Koller (2017). "A boy, being a victim, nobody really buys that, you know?": Dynamics of sexual violence against boys. *Child Abuse & Neglect 70*, 53–64.

Joseph, G., S. U. Javaid, L. A. Andres, J. Solotaroff, and S. I. Rajan (2017). Underreporting of

gender-based violence in Kerala, India : an application of the list randomization method. *World Bank Group WPS8044*.

Kramon, E. and K. Weghorst (2019). (mis) measuring sensitive attitudes with the list experiment: Solutions to list experiment breakdown in Kenya. *Public Opinion Quarterly 83*.

Lokot, M., A. Bhatia, L. Kenny, and B. Cislaghi (2020). Corporal punishment, discipline and social norms: A systematic review in low- and middle-income countries. *Aggression and Violent Behavior 55*, 101507.

López-Castro, L., P. K. Smith, S. Robinson, and A. Görzig (2023). Age differences in bullying victimisation and perpetration: Evidence from cross-cultural surveys. *Aggression and Violent Behavior 73*, 101888.

Malloy, L. C. and S. N. Stolzenberg (2019). Questioning Kids: Applying the Lessons from Developmentally-Sensitive Investigative Interviewing to the Research Context. *Journal of child psychology and psychiatry, and allied disciplines 60*(3), 325–327.

Nurjannah, N., R. S. Oktari, H. Nisa, W. Y. Viridanda, W. Aidina, and S.-J. Wang (2024). Urban children at risk of violence: A qualitative study of experiences of parents, teachers, and service providers of collaborative support. *Narra J 4*(2), e793.

Peterman, A., M. Dione, A. Le Port, J. Briaux, F. Lamesse, and M. Hidrobo (2025). Disclosure of Violence against Women and Girls in Senegal. *The World Bank Economic Review 39*(3), 614–631.

Punjabi, M., J. Norman, L. Edwards, and P. Muyingo (2021). Using ACASI to Measure Gender-Based Violence in Ugandan Primary Schools. *RTI Press RTI Press Research Brief*.

Scheithauer, H., T. Hayer, F. Petermann, and G. Jugert (2006). Physical, verbal, and relational forms of bullying among German students: age trends, gender differences, and correlates. *Aggressive Behavior 32*(3), 261–275. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ab.20128.

Smarrelli, G. (2024). Advancing the measurement of school-related violence among young children. Institution: American Economic Association.

Smarrelli, G. and Y. N. Wong (2025). School Violence Context in Low- and Middle-Income Countries.

Tanton, C., A. Bhatia, J. Pearlman, and K. Devries (2023). Increasing disclosure of school-related gender-based violence: lessons from a systematic review of data collection methods and existing survey research. *BMC Public Health 23*(1), 1012.

Tourangeau, R. and T. Yan (2007). Sensitive questions in surveys. *Psychological Bulletin 133*(5), 859–883.

Turchik, J. A. and K. M. Edwards (2012). Myths about male rape: A literature review. *Psychology of Men & Masculinity 13*(2), 211–226. Place: US Publisher: Educational Publishing Foundation.

UNESCO (2019). Behind the numbers: Ending school violence and bullying. Technical report, The United Nations Educational, Scientific and Cultural Organization.

von Russdorf, S., L. Ahlborn, A. Hidalgo-Arestegui, G. McQuade, and M. Favara (2024). A sound methodology: Measuring experiences of violent conflict through audio self-interviews. *Economics Letters 242*, 111879.

Zhou, Z., X. Zhou, G. Shen, A. Z. Khairani, and J. Saibon (2023). Correlates of Bullying Behavior Among Children and Adolescents in Physical Education: A Systematic Review. *Psychology Research and Behavior Management 16*, 5041–5051. Publisher: Dove Medical Press _eprint: https://www.tandfonline.com/doi/pdf/10.2147/PRBM.S441619.

A Table and Figures

Figure A.1: International Surveys with Questions on School-related Violence

Surveys	Administration mode	Location	Age	Bullying	Corporal Punishment	Sexual Violence
Demographic Health Surveys (DHS)	Face to Face	Home	15 to 49 (female only)		✓	✓
Multiple Indicator Cluster Survey (MICS)	Face to Face	Home	15 to 49 (caretakers)		✓	
Violence Against Children and Youth Surveys (VACS)	Face to Face	Home	13 to 24		✓	✓
Global School-based Student Health Survey (GSHS)	Self-administered	School	11 to 16	✓		
Health Behaviour in School-aged Children (HBSC)	Self-administered	School	11 to 15	✓		
Programme for International Student Assessment (PISA) Progress In International Reading Literacy Study (PIRLS) Estudio Regional Comparativo y Explicativo (ERCE) Trends In International Mathematics and Science Study (TIMSS)	Self-administered	School	15 / 9-12 / 10-14 / 10	✓		

Figure A.2: Primary Outcomes: Survey Questions

Emotional Bullying	In the last month, how often have other students from your school made fun of you or called you names.
	In the last month, how often have other students from your school left you out of their games or activities.
	In the last month, how often have other students from your school told lies or spread rumors about you.
	In the last month, how often have other students from your school threatened you.
Physical Bullying	In the last month, how often have other students hit, kicked, punched, or pushed you
	In the last month, how often have other students pinched you, pull your hair or pull your ears.
	In the last month, how often have other students thrown something at you to hurt you?
Emotional Violence by Staff	Think about the last month, how often did a teacher or member of the school staff embarrass, humiliate, or threaten you?
Corporal Punishment by Staff	Think about the last month, how often did a teacher or member of the school staff used physical punishment on you? Physical punishment includes spanking, beating, punching, twisting a child's ears or any other hitting, by using, for example, a hand or a stick.
Sexual Violence (Staff/Peers)	*At any point in your life, has X:*
Sexual Comments	Disturbed or bothered you by making comments, jokes, or gestures about parts of your body?
Nude photos/videos	Showed you nude or nearly nude pictures or videos that you didn't want to see or that made you feel uncomfortable?
Touch	Made you touch their body (for example, their bottom or other private parts) when you didn't want to or in a way that made you uncomfortable?
Touch	Touched your body (for example, touch your bottom, breast, penis, vagina, or other private parts) when you didn't want them to or in a way that made you uncomfortable?
Kiss	Kissed you when you didn't want to be kissed and in a way that made you uncomfortable?
Make the child take clothe off	Made you take your clothes off when you didn't want to and in a way that made you uncomfortable?
Forced sex	Made you have sex with them by threatening or pressuring you, or by making you afraid of what they might do?

Figure A.3: ACASI

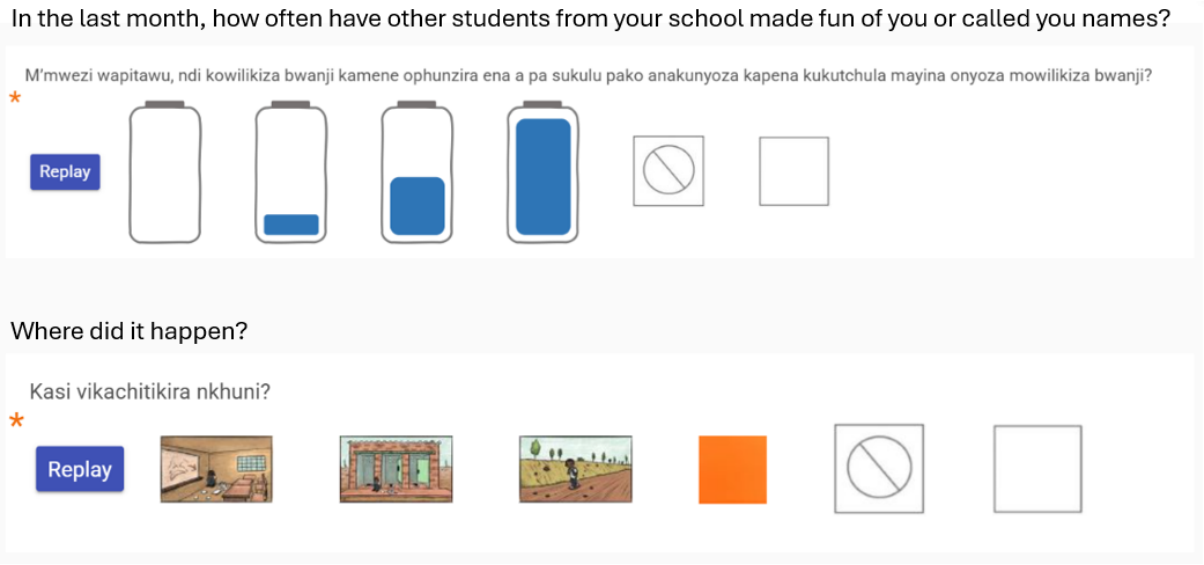

Table A.1: District characteristics

	All	*District without YONECO*	*District with YONECO*	*P-value*
Population in thousands	548.87	484.36	578.19	0.443
	(348.46)	(290.36)	(374.48)	
Log district population	6.04	5.81	6.15	0.416
	(0.90)	(1.23)	(0.71)	
Proportion of rural population	0.94	0.92	0.95	0.144
	(0.05)	(0.05)	(0.04)	
Primary: Enrolled students [in thousands]	165.58	145.13	174.87	0.403
	(102.82)	(82.81)	(111.25)	
Proportion of pop that is illiterate	0.29	0.28	0.30	0.657
	(0.11)	(0.09)	(0.11)	
Proportion of pop that never attended school	0.17	0.16	0.18	0.556
	(0.09)	(0.07)	(0.09)	
Proportion of unemployed pop	0.19	0.21	0.18	0.218
	(0.05)	(0.06)	(0.04)	
Children deprivation rate	0.58	0.57	0.58	0.851
	(0.18)	(0.19)	(0.18)	
Children monetary poverty rate	0.57	0.53	0.59	0.445
	(0.20)	(0.18)	(0.20)	
Health deprivation rate	0.04	0.05	0.04	0.189
	(0.02)	(0.03)	(0.02)	
Nutrition deprivation rate	0.31	0.28	0.32	0.288
	(0.11)	(0.10)	(0.11)	
Education deprivation rate	0.37	0.38	0.36	0.624
	(0.06)	(0.07)	(0.06)	
Water deprivation rate	0.14	0.17	0.12	0.123
	(0.07)	(0.10)	(0.05)	
Sanitation rate rate	0.28	0.24	0.30	0.048
	(0.09)	(0.07)	(0.10)	
Housing deprivation rate	0.50	0.51	0.50	0.963
	(0.19)	(0.22)	(0.18)	

Notes: The table presents means and standard deviations (in parentheses) between districts with and without YONECO presence. The last column shows the *p-values* from comparing the mean differences between the two groups of districts.

Table A.2: Sample characteristics

	(1) F2F Home	(2) F2F School	(3) ACASI Home	(4) ACASI School	(2)-(1)	(3)-(1)	(4)-(1)	(2)-(3)	(4)-(2)	(4)-(3)
Male Student	0.50	0.50	0.50	0.50	0.869	0.942	0.956	0.913	0.927	0.985
	(0.50)	(0.50)	(0.50)	(0.50)						
Studeng age	9.77	9.81	9.77	9.77	0.978	0.979	0.382	0.396	0.957	0.369
	(1.41)	(1.41)	(1.40)	(1.40)						
Old cohort [10 to 12]	0.51	0.52	0.51	0.51	0.987	0.826	0.812	0.825	0.814	0.648
	(0.50)	(0.50)	(0.50)	(0.50)						
Caregiver: complete secondary education or more	0.14	0.14	0.13	0.14	0.301	0.518	0.394	0.856	0.700	0.838
	(0.34)	(0.35)	(0.33)	(0.35)						
Wealth Index	0.39	0.40	0.39	0.40	0.161	0.788	0.254	0.788	0.095	0.158
	(0.20)	(0.20)	(0.20)	(0.20)						
School size	7.10	7.10	7.11	7.10	0.874	0.741	0.920	0.954	0.864	0.818
	(0.69)	(0.69)	(0.69)	(0.69)						

Notes: The table presents means and standard deviations (in parentheses) by treatment group. In the last six columns, we present the *p-values* from comparing the mean differences between the different treatment arms.

Table A.3: Differential Attrition

	(1)
ACASI Home	-0.001
	(0.002)
ACASI School	-0.001
	(0.002)
F2F School	-0.005
	(0.003)
Observations	6009

Notes: Regression coefficients show estimated average marginal effects. Robust standard errors are in parentheses. The dependent variable is a dummy variable that is equal to one if the respondent could not be surveyed. Statistical significance is indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Sexual Violence: Contact

	All			Staff			Peer		
	(1)	(2) Lower Bound	(3) Upper Bound	(4)	(5) Lower Bound	(6) Upper Bound	(7)	(8) Lower Bound	(9) Upper Bound
ACASI Home	0.092***	0.088***	0.109***	0.087***	0.075***	0.122***	0.061***	0.057***	0.085***
	(0.014)	(0.014)	(0.014)	(0.013)	(0.011)	(0.015)	(0.013)	(0.012)	(0.013)
ACASI School	0.110***	0.104***	0.132***	0.100***	0.087***	0.139***	0.073***	0.068***	0.102***
	(0.013)	(0.013)	(0.014)	(0.013)	(0.011)	(0.015)	(0.012)	(0.012)	(0.013)
F2F School	0.065***	0.063***	0.070***	0.024*	0.013	0.033**	0.058***	0.056***	0.066***
	(0.014)	(0.014)	(0.015)	(0.014)	(0.013)	(0.017)	(0.012)	(0.012)	(0.014)
Observations	5918	5982	5982	5895	5982	5982	5901	5982	5982
Mean	0.071	0.072	0.071	0.012	0.015	0.012	0.063	0.064	0.063
Pvalue: diff AH&AS	0.114	0.144	0.043	0.047	0.056	0.016	0.240	0.277	0.113
Pvalue: diff AS&FS	0.000	0.000	0.000	0.000	0.000	0.000	0.152	0.252	0.001
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group."Pvalue: diff AH&AS" refers to differences between the ACASI group that responded to the survey at school or home. "Pvalue: diff AS&FS" refers to differences between those who responded to the survey in school, using F2F or ACASI. Columns (1), (4), and (7) include all individual, family and school pre-specified covariates. Columns (2), (3), (5), (6), (8) and (9) show the Manski–Horowitz lower and upper bounds used to adjust for "don't know" and "decline" responses. Robust standard errors are shown in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Multiple Hypothesis Testing

	Emotional Bullying from Peers	Physical Bullying from Peers	Emotional Violence from School Staff	Corporal Punishment from School Staff	Sexual Violence from Peers	Sexual Violence from School Staff
ACASI Home						
p-value	0.000	0.001	0.000	0.000	0.000	0.000
RW - FWER	0.000	0.005	0.000	0.000	0.000	0.000
ACASI School						
p-value	0.000	0.000	0.000	0.000	0.000	0.000
RW - FWER	0.000	0.001	0.000	0.000	0.000	0.000
F2F School						
p-value	0.008	0.027	0.203	0.615	0.000	0.009
RW - FWER	0.044	0.120	0.356	0.616	0.000	0.044

Notes: The first row shows the robust standard errors in parentheses, while the second row, for each treatment arm, shows the standard errors computed using Romano-Wolf step-down adjusted p-values to control the family-wise error rate.

Table A.6: Sexual Violence: No Contact

	All			Staff			Peer		
	(1)	(2) Lower Bound	(3) Upper Bound	(4)	(5) Lower Bound	(6) Upper Bound	(7)	(8) Lower Bound	(9) Upper Bound
ACASI Home	0.108***	0.105***	0.118***	0.126***	0.120***	0.148***	0.053***	0.050***	0.071***
	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)	(0.015)	(0.014)	(0.013)	(0.014)
ACASI School	0.147***	0.143***	0.160***	0.146***	0.139***	0.169***	0.084***	0.080***	0.106***
	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)	(0.015)	(0.013)	(0.013)	(0.014)
F2F School	0.060***	0.060***	0.061***	0.034**	0.031**	0.038**	0.048***	0.047***	0.053***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.017)	(0.014)	(0.014)	(0.014)
Observations	5938	5982	5982	5917	5982	5982	5913	5982	5982
Mean	0.113	0.113	0.113	0.033	0.034	0.033	0.096	0.096	0.096
Pvalue: diff AH&AS	0.002	0.003	0.001	0.022	0.026	0.015	0.009	0.012	0.003
Pvalue: diff AS&FS	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.007	0.000
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group."Pvalue: diff AH&AS" refers to differences between the ACASI group that responded to the survey at school or home. "Pvalue: diff AS&FS" refers to differences between those who responded to the survey in school, using F2F or ACASI. Columns (1), (4), and (7) include all individual, family and school pre-specified covariates. Columns (2), (3), (5), (6), (8) and (9) show the Manski–Horowitz lower and upper bounds used to adjust for "don't know" and "decline" responses. Robust standard errors are shown in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Emotional Violence: Heterogeneity Analysis

	Individual Characteristics		Family Characteristics			Context-Level Characteristics	
	(1) Male student	(2) Old cohort	(3) Wealth	(4) Carer: secondary or more	(5) Christian Family	(6) School Size	(7) High Density City
ACASI*Male Student	-0.014						
	(0.024)						
ACASI*Old Cohort		0.005					
		(0.024)					
ACASI*Wealth Index			0.092				
			(0.060)				
ACASI*Carer has secondary education or more				0.033			
				(0.035)			
ACASI*Christian Family					0.027		
					(0.031)		
ACASI*(Log) School Size						0.028	
						(0.017)	
ACASI*High Density City							0.039
							(0.025)
Observations	5813	5813	5813	5813	5813	5813	5813

Notes: The table shows the estimated coefficients for the interaction terms between the ACASI treatment arm and individual, family, and context-level characteristics. The coefficients are estimated using equation 3.1, which includes interaction terms between the treatment variable and the covariates of interest, using Ordinary Least Squares. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Physical Violence: Heterogeneity Analysis

	Individual Characteristics		Family Characteristics			Context-Level Characteristics	
	(1) Male student	(2) Old cohort	(3) Wealth	(4) Carer: secondary or more	(5) Christian Family	(6) School Size	(7) High Density City
ACASI*Male Student	0.013						
	(0.024)						
ACASI*Old Cohort		0.048**					
		(0.024)					
ACASI*Wealth Index			-0.051				
			(0.061)				
ACASI*Carer has secondary education or more				0.067*			
				(0.036)			
ACASI*Christian Family					0.031		
					(0.032)		
ACASI*(Log) School Size						-0.011	
						(0.018)	
ACASI*High Density City							-0.013
							(0.025)
Observations	5828	5828	5828	5828	5828	5828	5828

Notes: The table shows the estimated coefficients for the interaction terms between the ACASI treatment arm and individual, family, and context-level characteristics. The coefficients are estimated using equation 3.1, which includes interaction terms between the treatment variable and the covariates of interest, using Ordinary Least Squares. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Sexual Violence: Heterogeneity Analysis

	Individual Characteristics		Family Characteristics			Context-Level Characteristics	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Male student	Old cohort	Wealth	Carer: secondary or more	Christian Family	School Size	High Density City
ACASI*Male Student	0.074***						
	(0.021)						
ACASI*Old Cohort		0.000					
		(0.021)					
ACASI*Wealth Index			0.099*				
			(0.053)				
ACASI*Carer has secondary education or more				0.053*			
				(0.030)			
ACASI*Christian Family					0.014		
					(0.026)		
ACASI*(Log) School Size						0.005	
						(0.015)	
ACASI*High Density City							-0.002
							(0.022)
Observations	5942	5942	5942	5942	5942	5942	5942

Notes: The table shows the estimated coefficients for the interaction terms between the ACASI treatment arm and individual, family, and context-level characteristics. The coefficients are estimated using equation 3.1, which includes interaction terms between the treatment variable and the covariates of interest, using Ordinary Least Squares. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: Correlates of Misreporting: Emotional Bullying by Peers

	Emotional Bullying				Method differences:p-values from tests of differences					
	F2F HOME	ACASI HOME	F2F SCHOOL	ACASI SCHOOL						
	(1)	(2)	(3)	(4)	(1) - (2)	(1) - (3)	(1) - (4)	(2) - (3)	(2) - (4)	(3) - (4)
Individual Characteristics										
Male student	0.056**	0.014	0.013	0.052**	0.237	0.227	0.917	0.978	0.276	0.264
	(0.025)	(0.025)	(0.025)	(0.025)						
Old cohort [10 to 12]	-0.021	-0.009	-0.069***	-0.013	0.748	0.168	0.821	0.085	0.923	0.104
	(0.025)	(0.025)	(0.025)	(0.025)						
Family Characteristics										
Wealth Index	-0.084	-0.033	-0.081	0.013	0.616	0.980	0.330	0.631	0.638	0.340
	(0.071)	(0.071)	(0.070)	(0.070)						
Caregiver: complete secondary or more	0.002	0.071*	-0.011	-0.010	0.222	0.818	0.839	0.133	0.138	0.977
	(0.041)	(0.039)	(0.039)	(0.038)						
Main practised religion: Christian	0.002	0.071*	-0.011	-0.010	0.674	0.894	0.184	0.776	0.359	0.236
	(0.041)	(0.039)	(0.039)	(0.038)						
Context-level: Characteristics										
High Density School	0.012	0.002	0.012	0.031	0.735	0.978	0.487	0.710	0.294	0.497
	(0.020)	(0.019)	(0.019)	(0.019)						
High Density City	-0.044	-0.002	-0.053*	-0.015	0.286	0.818	0.457	0.190	0.747	0.325
	(0.028)	(0.027)	(0.028)	(0.027)						

Notes: Robust standard errors are shown in parentheses. The last six columns show the two-sided p-value test difference between the treatment arms. p-values below 0.10 indicate that the characteristic is sensitive to the measurement method, providing suggestive evidence of systematic misreporting. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Correlates of Misreporting: Emotional Violence by the School Staff

	Emotional Violence				Method differences:p-values from tests of differences					
	F2F HOME (1)	ACASI HOME (2)	F2F SCHOOL (3)	ACASI SCHOOL (4)	(1) - (2)	(1) - (3)	(1) - (4)	(2) - (3)	(2) - (4)	(3) - (4)
Individual Characteristics										
Male student	0.079***	0.044*	0.070***	0.074***	0.299	0.771	0.885	0.446	0.410	0.903
	(0.022)	(0.026)	(0.022)	(0.026)						
Old cohort [10 to 12]	0.013	-0.029	0.004	-0.002	0.220	0.773	0.669	0.343	0.462	0.872
	(0.022)	(0.026)	(0.023)	(0.026)						
Family Characteristics										
Wealth Index	-0.037	-0.059	-0.071	-0.152**	0.815	0.702	0.228	0.907	0.371	0.400
	(0.061)	(0.073)	(0.063)	(0.073)						
Caregiver: complete secondary or more	-0.019	-0.011	-0.003	-0.003	0.884	0.744	0.756	0.882	0.880	0.989
	(0.035)	(0.041)	(0.034)	(0.041)						
Main practised religion: Christian	0.023	0.025	0.031	0.054	0.979	0.868	0.500	0.901	0.554	0.615
	(0.029)	(0.036)	(0.032)	(0.035)						
Context-level: Characteristics										
High Density School	-0.026	0.050**	-0.003	0.006	0.003	0.330	0.216	0.040	0.114	0.727
	(0.017)	(0.020)	(0.017)	(0.020)						
High Density City	-0.004	0.027	-0.042	0.025	0.414	0.289	0.452	0.076	0.950	0.086
	(0.025)	(0.029)	(0.026)	(0.029)						

Notes: Robust standard errors are shown in parentheses. The last six columns show the two-sided p-value test difference between the treatment arms. p-values below 0.10 indicate that the characteristic is sensitive to the measurement method, providing suggestive evidence of systematic misreporting. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: Correlates of Misreporting: Physical Bullying by Peers

	Physical Bullying				Method differences:p-values from tests of differences					
	F2F HOME (1)	ACASI HOME (2)	F2F SCHOOL (3)	ACASI SCHOOL (4)	(1) - (2)	(1) - (3)	(1) - (4)	(2) - (3)	(2) - (4)	(3) - (4)
Individual Characteristics										
Male student	0.049*	0.035	0.036	0.012	0.713	0.731	0.321	0.980	0.533	0.516
	(0.026)	(0.026)	(0.026)	(0.026)						
Old cohort [10 to 12]	-0.088***	-0.055**	-0.079***	-0.055**	0.357	0.791	0.358	0.510	0.999	0.512
	(0.026)	(0.026)	(0.026)	(0.026)						
Family Characteristics										
Wealth Index	0.022	-0.063	-0.181**	-0.012	0.403	0.045	0.736	0.250	0.623	0.101
	(0.071)	(0.073)	(0.073)	(0.073)						
Caregiver: complete secondary or more	-0.071*	0.032	0.016	0.053	0.074	0.126	0.031	0.774	0.728	0.520
	(0.041)	(0.041)	(0.040)	(0.041)						
Main practised religion: Christian	-0.013	0.100***	0.050	0.037	0.024	0.209	0.315	0.326	0.213	0.798
	(0.035)	(0.036)	(0.036)	(0.036)						
Context-level: Characteristics										
High Density School	-0.009	0.024	0.007	-0.019	0.231	0.553	0.745	0.541	0.130	0.359
	(0.020)	(0.020)	(0.020)	(0.020)						
High Density City	0.007	-0.004	-0.027	-0.021	0.790	0.414	0.492	0.583	0.675	0.897
	(0.029)	(0.029)	(0.029)	(0.029)						

Notes: Robust standard errors are shown in parentheses. The last six columns show the two-sided p-value test difference between the treatment arms. p-values below 0.10 indicate that the characteristic is sensitive to the measurement method, providing suggestive evidence of systematic misreporting. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.13: Correlates of Misreporting: Corporal Punishment by the School Staff

	Corporal Punishment				Method differences:p-values from tests of differences					
	F2F HOME (1)	ACASI HOME (2)	F2F SCHOOL (3)	ACASI SCHOOL (4)	(1) - (2)	(1) - (3)	(1) - (4)	(2) - (3)	(2) - (4)	(3) - (4)
Individual Characteristics										
Male student	0.052**	0.010	-0.002	0.082***	0.244	0.137	0.420	0.750	0.050	0.022
	(0.026)	(0.026)	(0.026)	(0.026)						
Old cohort [10 to 12]	-0.010	0.013	0.005	0.021	0.528	0.681	0.391	0.825	0.822	0.655
	(0.026)	(0.026)	(0.026)	(0.026)						
Family Characteristics										
Wealth Index	0.082	-0.066	0.103	-0.085	0.149	0.841	0.103	0.102	0.852	0.068
	(0.072)	(0.073)	(0.072)	(0.073)						
Caregiver: complete secondary or more	-0.135***	0.025	0.008	0.013	0.006	0.013	0.011	0.763	0.830	0.930
	(0.041)	(0.042)	(0.040)	(0.041)						
Main practised religion: Christian	0.020	0.075**	0.080**	0.020	0.280	0.241	0.997	0.923	0.283	0.244
	(0.035)	(0.036)	(0.036)	(0.036)						
Context-level: Characteristics										
High Density School	-0.046**	-0.017	-0.021	-0.051**	0.298	0.377	0.859	0.872	0.224	0.289
	(0.020)	(0.020)	(0.020)	(0.020)						
High Density City	-0.017	-0.051*	-0.072**	-0.024	0.413	0.180	0.865	0.596	0.516	0.241
	(0.029)	(0.029)	(0.029)	(0.029)						

Notes: Robust standard errors are shown in parentheses. The last six columns show the two-sided p-value test difference between the treatment arms. p-values below 0.10 indicate that the characteristic is sensitive to the measurement method, providing suggestive evidence of systematic misreporting. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.14: Correlates of Misreporting: Sexual Violence by Peers

	Sexual Violence by Peers				Method differences:p-values from tests of differences					
	F2F HOME (1)	ACASI HOME (2)	F2F SCHOOL (3)	ACASI SCHOOL (4)	(1) - (2)	(1) - (3)	(1) - (4)	(2) - (3)	(2) - (4)	(3) - (4)
Individual Characteristics										
Male student	0.021	0.077***	0.011	0.047**	0.034	0.729	0.345	0.022	0.307	0.233
	(0.017)	(0.020)	(0.020)	(0.022)						
Old cohort [10 to 12]	0.039**	0.053***	0.073***	0.057***	0.581	0.186	0.497	0.485	0.894	0.586
	(0.017)	(0.020)	(0.020)	(0.022)						
Family Characteristics										
Wealth Index	-0.083*	0.137**	0.046	0.025	0.003	0.076	0.158	0.269	0.189	0.802
	(0.044)	(0.059)	(0.058)	(0.062)						
Caregiver: complete secondary or more	-0.031	-0.054*	-0.029	0.023	0.571	0.941	0.195	0.566	0.104	0.265
	(0.024)	(0.032)	(0.031)	(0.035)						
Main practised religion: Christian	-0.031	-0.054*	-0.029	0.023	0.933	0.393	0.812	0.395	0.889	0.326
	(0.024)	(0.032)	(0.031)	(0.035)						
Context-level: Characteristics										
High Density School	-0.016	0.008	0.003	-0.017	0.237	0.320	0.963	0.811	0.273	0.360
	(0.012)	(0.017)	(0.015)	(0.016)						
High Density City	0.036*	0.012	0.028	0.030	0.424	0.791	0.842	0.619	0.584	0.954
	(0.020)	(0.023)	(0.023)	(0.023)						

Notes: Robust standard errors are shown in parentheses. The last six columns show the two-sided p-value test difference between the treatment arms. p-values below 0.10 indicate that the characteristic is sensitive to the measurement method, providing suggestive evidence of systematic misreporting. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.15: Correlates of Misreporting: Sexual Violence by the School Staff

	Sexual Violence by School Staff				Method differences:p-values from tests of differences					
	F2F HOME (1)	ACASI HOME (2)	F2F SCHOOL (3)	ACASI SCHOOL (4)	(1) - (2)	(1) - (3)	(1) - (4)	(2) - (3)	(2) - (4)	(3) - (4)
Individual Characteristics										
Male student	0.008	0.059***	0.029**	0.083***	0.018	0.188	0.001	0.180	0.383	0.021
	(0.010)	(0.019)	(0.012)	(0.021)						
Old cohort [10 to 12]	0.014	0.004	0.013	0.036*	0.646	0.977	0.331	0.678	0.255	0.344
	(0.010)	(0.019)	(0.012)	(0.021)						
Family Characteristics										
Wealth Index	-0.020	0.050	-0.009	-0.054	0.220	0.767	0.594	0.337	0.187	0.497
	(0.024)	(0.052)	(0.032)	(0.059)						
Caregiver: complete secondary or more	-0.004	-0.031	-0.018	0.003	0.421	0.512	0.845	0.715	0.445	0.560
	(0.014)	(0.030)	(0.017)	(0.032)						
Main practised religion: Christian	0.009	-0.031	0.016	-0.003	0.186	0.736	0.695	0.143	0.466	0.555
	(0.013)	(0.027)	(0.017)	(0.027)						
Context-level: Characteristics										
High Density School	0.002	-0.011	-0.004	-0.004	0.422	0.607	0.712	0.684	0.744	0.992
	(0.007)	(0.015)	(0.010)	(0.016)						
High Density City	-0.002	0.038*	-0.003	0.028	0.105	0.956	0.236	0.120	0.757	0.251
	(0.012)	(0.022)	(0.015)	(0.023)						

Notes: Robust standard errors are shown in parentheses. The last six columns show the two-sided p-value test difference between the treatment arms. p-values below 0.10 indicate that the characteristic is sensitive to the measurement method, providing suggestive evidence of systematic misreporting. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.4: Safeguarding protocol

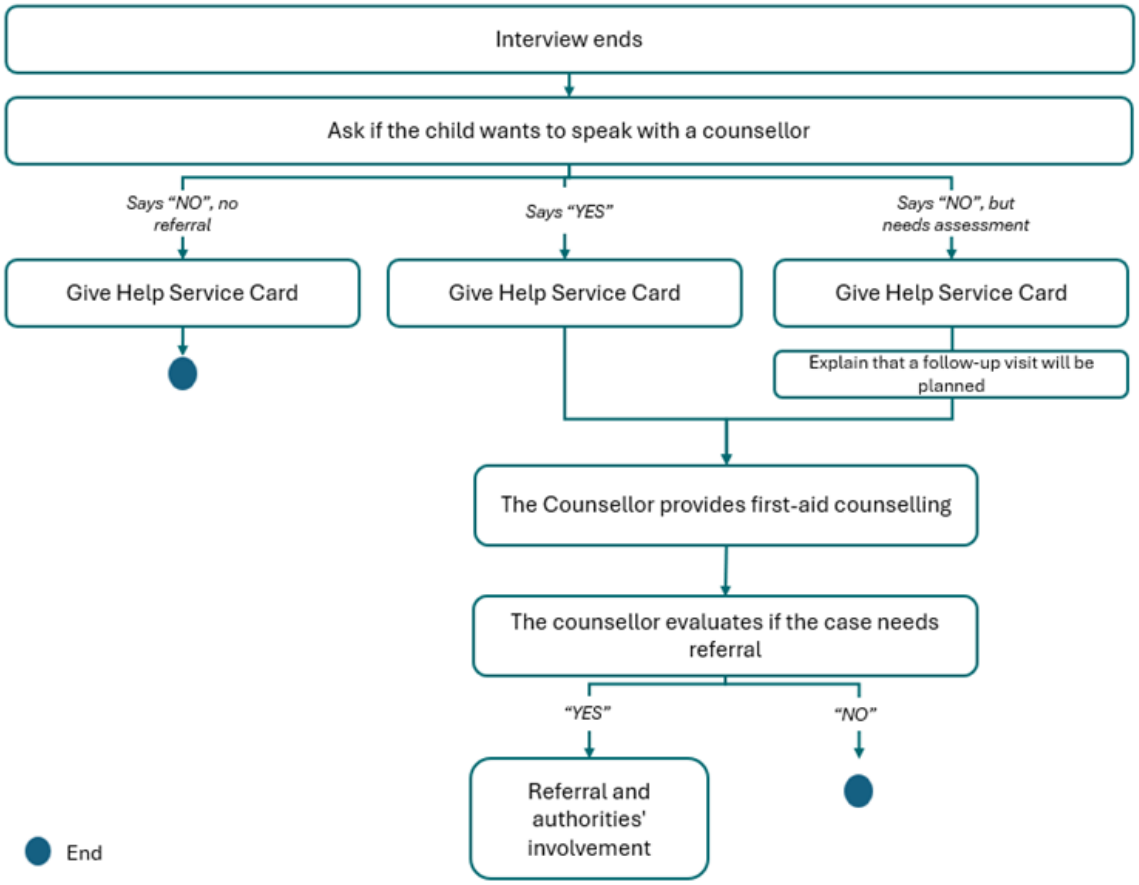

Table A.16: Mean Differences by Attention Check Outcome

	All	*Failed attention check*	*Didn't fail attention check*	*P-value*
Male Student	0.50	0.56	0.49	0.002
	(0.50)	(0.50)	(0.50)	
Age	9.78	9.44	9.82	0.000
	(1.41)	(1.30)	(1.41)	
Old Cohort	0.51	0.43	0.53	0.000
	(0.50)	(0.49)	(0.50)	
Caregiver has secondary education	0.14	0.10	0.14	0.000
	(0.34)	(0.29)	(0.35)	
Wealth Index	0.40	0.37	0.40	0.000
	(0.20)	(0.19)	(0.20)	

Notes: The table presents means and standard deviations (in parentheses) by those that failed and did not failed the attention check questions. The last column shows the two-sided p-values test difference between those that failed and did not failed the attention check questions. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.17: Dropping Observations that Failed Attention Checks

	Physical Bullying		Emotional Bullying		Emotional Violence Staff		Corporal P. Staff		Sexual V Staff		Sexual V Peer	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ACASI Home	0.095***	0.087***	0.059***	0.058***	0.206***	0.208***	0.088***	0.078***	0.151***	0.142***	0.076***	0.081***
	(0.018)	(0.019)	(0.017)	(0.018)	(0.017)	(0.017)	(0.018)	(0.019)	(0.015)	(0.015)	(0.015)	(0.016)
ACASI School	0.085***	0.088***	0.065***	0.064***	0.208***	0.205***	0.093***	0.090***	0.173***	0.159***	0.101***	0.090***
	(0.018)	(0.019)	(0.017)	(0.019)	(0.017)	(0.017)	(0.018)	(0.019)	(0.015)	(0.015)	(0.015)	(0.016)
F2F School	0.048***	0.054***	0.038**	0.041**	0.023	0.024	-0.008	-0.009	0.044***	0.040**	0.067***	0.070***
	(0.018)	(0.019)	(0.017)	(0.018)	(0.018)	(0.019)	(0.018)	(0.019)	(0.017)	(0.017)	(0.015)	(0.016)
Observations	5896	5249	5888	5243	5873	5236	5876	5237	5925	5273	5922	5272
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group. Columns (2), (4), (6), (8), (10) and (12) drop all observations that failed the attention checks. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.18: Controlling for Attention Check Failure

	Physical Bullying		Emotional Bullying		Emotional Violence Staff		Corporal P. Staff		Sexual V Staff		Sexual V Peer	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ACASI Home	0.095***	0.092***	0.059***	0.057***	0.206***	0.204***	0.088***	0.087***	0.151***	0.150***	0.076***	0.076***
	(0.016)	(0.016)	(0.018)	(0.018)	(0.016)	(0.016)	(0.019)	(0.019)	(0.015)	(0.015)	(0.014)	(0.014)
ACASI School	0.085***	0.081***	0.065***	0.062***	0.208***	0.204***	0.093***	0.091***	0.173***	0.170***	0.101***	0.100***
	(0.018)	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.014)	(0.014)	(0.014)	(0.014)
F2F School	0.048***	0.049***	0.038**	0.038**	0.023	0.024	-0.008	-0.008	0.044***	0.044***	0.067***	0.067***
	(0.017)	(0.017)	(0.016)	(0.016)	(0.017)	(0.017)	(0.019)	(0.019)	(0.016)	(0.016)	(0.014)	(0.014)
Observations	5896	5896	5888	5888	5873	5873	5876	5876	5925	5925	5922	5922
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control for Attention Check		Yes		Yes		Yes		Yes		Yes		Yes

Notes: The table presents estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group. Columns (2), (4), (6), (8), (10) and (12) control for a dummy variable that takes the value of one if the individual failed the attention checks. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.19: Evidence of Straightlining

	Dummy	Raw Total
	(1)	(2)
ACASI Home	-0.022*	0.243***
	(0.013)	(0.039)
ACASI School	-0.014	0.251***
	(0.013)	(0.038)
F2F School	-0.046***	-0.118***
	(0.013)	(0.035)
Observations	5982	5982
mean	1.499	1.499
Controls	Yes	

Notes: The outcome used in the first column is a dummy variable that takes the value 1 if straightlining occurs in at least one module and 0 otherwise. The outcome used in the second column is the total number of modules with suggestive evidence of straightlining. Robust standard errors are shown in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.20: Peer to Peer Violence: Enumerator Fixed Effects

	Emotional Bullying		Physical Bullying		Sexual Violence	
	(1)	(2)	(3)	(4)	(5)	(6)
ACASI Home	0.059***	0.060***	0.095***	0.095***	0.076***	0.081***
	(0.017)	(0.017)	(0.018)	(0.017)	(0.015)	(0.014)
ACASI School	0.065***	0.062***	0.085***	0.084***	0.101***	0.103***
	(0.017)	(0.017)	(0.018)	(0.018)	(0.015)	(0.014)
F2F School	0.038**	0.037**	0.048***	0.047***	0.067***	0.070***
	(0.017)	(0.017)	(0.018)	(0.017)	(0.015)	(0.014)
Observations	5888	5887	5896	5888	5922	5907
Mean	0.608		0.456		0.124	
Pvalue: diff AH&AS	0.745		0.622		0.063	
Pvalue: diff AS&FS	0.124		0.042		0.010	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Enumerator FE	No	Yes	No	Yes	No	Yes

Notes: Columns (1), (3) and (4) show the estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group. Columns (2), (4) and (6) include enumerator fixed effects.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.21: School Staff to Student Violence: Enumerator Fixed Effects

	Emotional Bullying		Physical Bullying		Sexual Violence	
	(1)	(2)	(3)	(4)	(5)	(6)
ACASI Home	0.206***	0.205***	0.088***	0.088***	0.151***	0.158***
	(0.017)	(0.016)	(0.018)	(0.017)	(0.015)	(0.014)
ACASI School	0.208***	0.207***	0.093***	0.091***	0.173***	0.179***
	(0.017)	(0.016)	(0.018)	(0.017)	(0.015)	(0.014)
F2F School	0.023	0.024	-0.009	-0.009	0.043***	0.046***
	(0.018)	(0.017)	(0.018)	(0.017)	(0.017)	(0.016)
Observations	5873	5858	5876	5869	5925	5659
Mean	0.233		0.473		0.038	
Pvalue: diff AH&AS	0.903		0.775		0.019	
Pvalue: diff AS&FS	0.000		0.000		0.000	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Enumerator FE	No	Yes	No	Yes	No	Yes

Notes: Columns (1), (3) and (4) show the estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group. Columns (2), (4) and (6) include enumerator fixed effects.

Table A.22: Mean Differences Between the Full Sample and the Counselling Sample

	All	***Counselling Provided***	***Counselling Not Provided***	***P-value***
Child Sex	0.50	0.54	0.48	0.000
	(0.50)	(0.50)	(0.50)	
Child Age	9.78	9.96	9.72	0.000
	(1.41)	(1.40)	(1.40)	
1: Old cohort [10 to 12]	0.51	0.57	0.49	0.000
	(0.50)	(0.49)	(0.50)	
Caregiver: complete secondary or more	0.14	0.14	0.14	0.890
	(0.34)	(0.34)	(0.34)	
Wealth Index	0.40	0.41	0.39	0.011
	(0.20)	(0.21)	(0.20)	

Notes: The table presents means and standard deviations (in parentheses) for those in the ACASI or F2F group among children who were referred for counselling. The last column shows the two-sided p-values. Robust standard errors are shown in parentheses.

Table A.23: Mean Differences by Evidence of Question Misunderstanding

	All	*Evidence of misunderstanding*	*No evidence of misunderstanding*	*P-value*
Child Sex	0.54	0.66	0.53	0.001
	(0.50)	(0.48)	(0.50)	
Child Age	9.96	9.82	9.97	0.190
	(1.40)	(1.37)	(1.40)	
1: Old cohort [10 to 12]	0.57	0.51	0.58	0.114
	(0.49)	(0.50)	(0.49)	
Caregiver: complete secondary or more	0.14	0.12	0.14	0.456
	(0.34)	(0.32)	(0.34)	
Wealth Index	0.41	0.39	0.41	0.192
	(0.21)	(0.20)	(0.21)	

Notes: The table presents means and standard deviations (in parentheses) for those with evidence of misunderstanding a question, and those with no evidence of misunderstanding a question. The last column shows the two-sided p-values. Robust standard errors are shown in parentheses.

Table A.24: Balance by Treatment Arm in the Counselling Sample

	All	***ACASI***	***F2F***	***P-value***
Child Sex	0.54	0.55	0.52	0.163
	(0.50)	(0.50)	(0.50)	
Child Age	9.96	9.95	9.96	0.939
	(1.40)	(1.41)	(1.38)	
1: Old cohort [10 to 12]	0.57	0.57	0.58	0.449
	(0.49)	(0.50)	(0.49)	
Caregiver: complete secondary or more	0.14	0.13	0.14	0.819
	(0.34)	(0.34)	(0.34)	
Wealth Index	0.41	0.41	0.40	0.357
	(0.21)	(0.20)	(0.21)	

Notes: The table presents means and standard deviations (in parentheses) for those in the ACASI or F2F group among children who were referred for counselling. The last column shows the two-sided p-values. Robust standard errors are shown in parentheses.

Table A.25: Correlates of the Likelihood of Misunderstanding Survey Questions

	(1)	(2)	(3)	(4)	(5)	(6)	est7
1=ACASI interview	0.061***	0.045*	0.067**	0.067***	0.046	0.061	0.058
	(0.014)	(0.019)	(0.024)	(0.016)	(0.033)	(0.135)	(0.139)
Child Sex	0.047**	0.028					0.029
	(0.015)	(0.018)					(0.018)
1: Old cohort [10 to 12]	-0.023		-0.019				-0.018
	(0.015)		(0.019)				(0.019)
Wealth Index	-0.032				-0.073		-0.072
	(0.041)				(0.046)		(0.054)
Caregiver: complete secondary or more	-0.011			0.001			0.017
	(0.023)			(0.027)			(0.029)
Log: Total enrolled students	-0.016					-0.019	-0.014
	(0.011)					(0.012)	(0.013)
ACAS*Child Sex		0.030					0.033
		(0.029)					(0.029)
ACAS*1: Old cohort [10 to 12]			-0.008				-0.010
			(0.030)				(0.030)
ACAS*Caregiver: complete secondary or more				-0.029			-0.051
				(0.041)			(0.045)
ACAS*Wealth Index					0.042		0.072
					(0.070)		(0.080)
ACAS*Log: Total enrolled students						0.000	-0.004
						(0.019)	(0.020)
Constant	0.184*	0.049***	0.074***	0.063***	0.092***	0.196*	0.182*
	(0.072)	(0.012)	(0.016)	(0.010)	(0.023)	(0.088)	(0.092)
N	1566.000	1566.000	1566.000	1566.000	1566.000	1566.000	1566.000
Mean	0.075						

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Column (1) includes individual and household controls. Columns (2) to(5) include individual and household controls one at a time, plus an interaction between the control and the ACASI treatment. Column (6) includes all controls and interactions. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.26: Correlates of Misreporting: Sexual Violence Adjusted for Potential Misunderstandings

	Sexual Violence		Diff (p-value)	Sexual Violence (Adj)		Diff (p-value)
	ACASI	F2F		ACASI	F2F	
	(1)	(2)	(1) - (2)	(1)	(2)	(1) - (2)
Individual Characteristics						
Male student	0.095***	0.024*	0.001	0.069***	0.018	0.013
	(0.016)	(0.014)		(0.016)	(0.013)	
Old cohort [10 to 12]	0.058***	0.060***	0.934	0.058***	0.057***	0.960
	(0.016)	(0.014)		(0.016)	(0.013)	
Family Characteristics						
Wealth Index	0.060	-0.020	0.186	0.036	-0.013	0.396
	(0.047)	(0.038)		(0.045)	(0.037)	
Caregiver: complete secondary or more	-0.014	-0.042**	0.389	0.001	-0.044**	0.164
	(0.026)	(0.021)		(0.025)	(0.020)	
Main practised religion: Christian	0.045**	0.029	0.578	0.047**	0.015	0.233
	(0.022)	(0.019)		(0.020)	(0.018)	
Environment: Characteristics						
High Density School	-0.008	-0.007	0.943	-0.004	-0.006	0.905
	(0.013)	(0.010)		(0.012)	(0.010)	
High Density City	0.016	0.029*	0.578	0.036**	0.043***	0.780
	(0.018)	(0.016)		(0.017)	(0.015)	

Notes: The Table shows the relationship between child, family, and context-level characteristics and the likelihood of sexual violence by survey method, differentiating by the sexual violence reported in the survey and the one adjusted for potential misunderstandings. The Columns titled 'Diff (pvalue)' show the two-sided p-value test difference between the ACASI and F2F groups. p-values below 0.10 indicate that the characteristic is sensitive to the measurement method, providing suggestive evidence of systematic misreporting. Robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B PDS LASSO

The treatment effects shown in Tables 1, 2, and 3 include controls selected using post-double-selection LASSO (PDSL) (Belloni et al. 2014). The results remain similar to those obtained with our prespecified controls. For each regression, PDS Lasso picked different controls, including some that we pre-specified, such as child sex and age, and new ones, including interaction terms between age and gender.

Table A.27: Emotional and Physical Bullying by Peers

	Emotional Bullying		Physical Bullying	
	(1)	(2)	(3)	(4)
	Logit	PDSLasso	Logit	PDSLasso
ACASI Home	0.059***	0.060***	0.095***	0.095***
	(0.017)	(0.017)	(0.018)	(0.018)
ACASI School	0.065***	0.065***	0.085***	0.085***
	(0.017)	(0.017)	(0.018)	(0.018)
F2F School	0.038**	0.039**	0.048***	0.047***
	(0.017)	(0.017)	(0.018)	(0.018)
Observations	5888	5888	5896	5896
Controls	Yes	Yes	Yes	Yes

Notes: Columns (1) and (3) show the estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group. Columns (2) and (4) include controls selected using PDS Lasso. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.28: Emotional Violence and Corporal Punishment by the School Staff

	Emotional Violence		Corporal Punishment	
	(1)	(2)	(3)	(4)
	Logit	PDSLasso	Logit	PDSLasso
ACASI Home	0.206***	0.210***	0.088***	0.089***
	(0.017)	(0.017)	(0.018)	(0.018)
ACASI School	0.208***	0.210***	0.093***	0.092***
	(0.017)	(0.017)	(0.018)	(0.018)
F2F School	0.023	0.019	-0.009	-0.009
	(0.018)	(0.017)	(0.018)	(0.018)
Observations	5873	5873	5876	5876
Controls	Yes	Yes	Yes	Yes

Notes: Columns (1) and (3) show the estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group. Columns (2) and (4) include controls selected using PDS Lasso.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.29: Sexual Violence

	Sexual Violence by Peers		Sexual violence by School Staff	
	(1)	(2)	(3)	(4)
	Logit	PDSLasso	Logit	PDSLasso
ACASI Home	0.076***	0.068***	0.151***	0.125***
	(0.015)	(0.014)	(0.015)	(0.011)
ACASI School	0.101***	0.096***	0.173***	0.156***
	(0.015)	(0.014)	(0.015)	(0.011)
F2F School	0.067***	0.058***	0.043***	0.020*
	(0.015)	(0.014)	(0.017)	(0.011)
Observations	5922	5922	5925	5925
Controls	Yes	Yes	Yes	Yes

Notes: Columns (1) and (3) show the estimated average marginal effects of the treatment arms relative to the children allocated to the F2F Home group. Columns (2) and (4) include controls selected using PDS Lasso.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C Cost effectiveness analysis

Table A.30: Cost-effectiveness Ratio: Disclosure of Sexual Violence

Treatment arm	Cost per beneficiary	Indicator	Percentage point change compared to the control group	Cost-effectiveness ratio	Interpretation
F2F-school	60	Change in reporting of sexual violence	7.8	**7.69**	1 percentage point increase in the probability of reporting violence costs 7.69 USD when collecting data using F2F-school.
ACASI-home	65		12.8	**5.10**	1 percentage point increase in the probability of reporting violence costs 5.10 USD when collecting data using ACASI-home.
ACASI-school	61		16.7	**3.67**	1 percentage point increase in the probability of reporting violence costs 3.67 USD when collecting data using ACASI-school.

Notes: The cost-effectiveness ratio is calculated as the cost per beneficiary divided by the treatment effect, where the treatment effect is measured as the percentage-point change in violence disclosure. The calculation is based on the sexual violence outcome.

Table A.31: Cost-effectiveness Ratio: Disclosure of Sexual Violence (Adjusted)

Treatment arm	Cost per beneficiary	Indicator	Percentage point change compared to the control group	Cost-effectiveness ratio	Interpretation
F2F-school	60	Change in reporting of sexual violence	7.1	**8.44**	1 percentage point increase in the probability of reporting violence costs 8.44 USD when collecting data using F2F-school.
ACASI-home	65		10.2	**6.40**	1 percentage point increase in the probability of reporting violence costs 6.40 USD when collecting data using ACASI-home.
ACASI-school	61		13.8	**4.44**	1 percentage point increase in the probability of reporting violence costs 4.44 USD when collecting data using ACASI-school.

Notes: The cost-effectiveness ratio is calculated as the cost per beneficiary divided by the treatment effect, where the treatment effect is measured as the percentage-point change in violence disclosure. The calculation is based on the sexual violence outcome, adjusting for potential misunderstandings.

D Comparison with International Surveys

Comparing our estimates with those from international surveys involves several caveats. Ideally, we would compare our results with surveys that target the same age group, use identical survey questions, and have been administered in Malawi. Malawi has data from the GSHS (2009), the VACS (2013), and the DHS (2024). We exclude the DHS from the comparison because it only provides data on sexual violence for females aged 15 and older. Between the VACS and the GSHS, we prioritize comparability in measuring sexual violence and therefore focus on the VACS, as the GSHS does not collect such information.

The VACS is administered to boys and girls aged 13 to 24 using face-to-face interviews in respondents' homes and asks about violence experienced in any setting and by any perpetrator. We restrict the analysis to children aged 13 to 17. The survey includes two questions on forced or pressured sex: "Has anyone ever physically forced you to have sex and did succeed?" and "Has anyone ever pressured you to have sex, through harassment, threats, or tricks and did succeed?" It also includes a question on attempted sex: "Has anyone ever tried to make you have sex against your will but did not succeed?"

Three differences between the VACS and our survey are important. First, the VACS covers an older age group, so the degree of misreporting may differ across surveys. If victims increasingly internalize norms that discourage disclosure with age, our estimates of misreporting may be conservative. Second, our survey question differs somewhat from the VACS question. We ask: "At any point in your life, did a peer or member of the school staff make you have sex with them by threatening or pressuring you, or by making you afraid of what they might do?" Our measure focuses on forced and pressured sex and does not capture attempted sex. Although the wording differs, the underlying concepts are broadly comparable. Third, our survey focuses exclusively on school-related violence. We identify whether the perpetrator was a peer or school staff member for all cases, whereas the VACS records perpetrator information only for the first and most recent incident. As a result, for children experiencing more than two incidents, the VACS does not allow full identification of perpetrators. Based on the VACS data, this limitation affects approximately 18 percent of respondents.

These caveats imply that any extrapolation to existing surveys should be treated with considerable caution and viewed only as an illustrative exercise. Under the assumption that the level of misreporting estimated in our study applies to the VACS measure of forced or pressured

sex, the implied prevalence would be seven times higher than indicated by surveys such as the Malawi VACS. A similar exercise across 15 Sub-Saharan African countries with available VACS data would imply prevalence estimates exceeding those reported in the VACS by a factor of ten. If we instead include the VACS measure of attempted sex, the implied prevalence would be approximately twice as high as current survey estimates suggest, and exceeding reported rates by a factor of five if we consider the sample of 15 Sub-Saharan African countries. This latter comparison should be interpreted particularly cautiously, as our measure primarily captures forced and pressured sex rather than attempted sex, and would only be appropriate if some children answered the question referring to attempted sex instead of forced or pressured sex. We view this as unlikely, given that our estimates are based on measures already adjusted for potential misunderstandings.